The Handbook of Phonological Theory

Second Edition

Edited by

John Goldsmith, Jason Riggle, and Alan C. L. Yu

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

The Handbook of Phonological Theory

Blackwell Handbooks in Linguistics

This outstanding multi-volume series covers all the major subdisciplines within linguistics today and, when complete, will offer a comprehensive survey of linguistics as a whole.

Already published:

The Handbook of Child Language
Edited by Paul Fletcher and Brian MacWhinney

The Handbook of Phonological Theory
Edited by John A. Goldsmith

The Handbook of Contemporary Semantic Theory
Edited by Shalom Lappin

The Handbook of Sociolinguistics
Edited by Florian Coulmas

The Handbook of Phonetic Sciences, Second Edition
Edited by William J. Hardcastle and John Laver

The Handbook of Morphology
Edited by Andrew Spencer and Arnold Zwicky

The Handbook of Japanese Linguistics
Edited by Natsuko Tsujimura

The Handbook of Linguistics
Edited by Mark Aronoff and Janie Rees-Miller

The Handbook of Contemporary Syntactic Theory
Edited by Mark Baltin and Chris Collins

The Handbook of Discourse Analysis
Edited by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton

The Handbook of Language Variation and Change
Edited by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes

The Handbook of Historical Linguistics
Edited by Brian D. Joseph and Richard D. Janda

The Handbook of Language and Gender
Edited by Janet Holmes and Miriam Meyerhoff

The Handbook of Second Language Acquisition
Edited by Catherine J. Doughty and Michael H. Long

The Handbook of Bilingualism
Edited by Tej K. Bhatia and William C. Ritchie

The Handbook of Pragmatics
Edited by Laurence R. Horn and Gregory Ward

The Handbook of Applied Linguistics
Edited by Alan Davies and Catherine Elder

The Handbook of Speech Perception
Edited by David B. Pisoni and Robert E. Remez

The Blackwell Companion to Syntax, Volumes I–V
Edited by Martin Everaert and Henk van Riemsdijk

The Handbook of the History of English
Edited by Ans van Kemenade and Bettelou Los

The Handbook of English Linguistics
Edited by Bas Aarts and April McMahon

The Handbook of World Englishes
Edited by Braj B. Kachru; Yamuna Kachru, and Cecil L. Nelson

The Handbook of Educational Linguistics
Edited by Bernard Spolsky and Francis M. Hult

The Handbook of Clinical Linguistics
Edited by Martin J. Ball, Michael R. Perkins, Nicole Müller, and Sara Howard

The Handbook of Pidgin and Creole Studies
Edited by Silvia Kouwenberg and John Victor Singler

The Handbook of Language Teaching
Edited by Michael H. Long and Catherine J. Doughty

The Handbook of Language Contact
Edited by Raymond Hickey

The Handbook of Language and Speech Disorders
Edited by Jack S. Damico, Nicole Müller, Martin J. Ball

The Handbook of Clinical Linguistics and Natural Language Processing
Edited by Alexander Clark, Chris Fox, and Shalom Lappin

The Handbook of Language and Globalization
Edited by Nikolas Coupland

The Handbook of Hispanic Linguistics
Edited by Manuel Díaz-Campos

The Handbook of Phonological Theory

Second Edition

Edited by

John Goldsmith, Jason Riggle, and Alan C. L. Yu

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2011
© 2011 Blackwell Publishing Ltd.

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of John Goldsmith, Jason Riggle, and Alan C. L. Yu to be identified as the authors of the editorial material in this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data
The handbook of phonological theory / edited by John Goldsmith, Jason Riggle, and Alan Yu. – 2nd ed.
p. cm. – (Blackwell handbooks in linguistics)
Includes bibliographical references and index.
ISBN 978-1-4051-5768-1 (alk. paper)
1. Grammar, Comparative and general–Phonology. I. Goldsmith, John A., 1951– II. Riggle, Jason. III. Yu, Alan.
P217.H27 2011
414–dc22

2011008798

A catalogue record for this book is available from the British Library.

This book is published in the following electronic formats: ePDFs [ISBN 9781444343038]; Wiley Online Library [ISBN 9781444343069]

Set in 10/12pt Palatino by Graphicraft Limited, Hong Kong

1 2011

Contents

List of Contributors vii
Preface ix

1 Rules v. Constraints 1
David Odden

2 Opacity and Ordering 40
Eric Baković

3 The Interaction Between Morphology and Phonology 68
Sharon Inkelas

4 Quantity 103
Stuart Davis

5 Stress Systems 141
Matthew Gordon

6 The Syllable 164
John Goldsmith

7 Tone: Is it Different? 197
Larry M. Hyman

8 Harmony Systems 240
Sharon Rose and Rachel Walker

9 Contrast Reduction 291
Alan C. L. Yu

10 Diachronic Explanations of Sound Patterns 319
Gunnar Ólafur Hansson

11 Phonetics in Phonology 348
D. R. Ladd

12 Corpora and Exemplars in Phonology 374
Mirjam Ernestus and R. Harald Baayen

13 The Place of Variation in Phonological Theory 401
Andries W. Coetzee and Joe Pater

14 The Syntax-Phonology Interface 435
Elisabeth Selkirk

15 Intonation 485
Mary E. Beckman and Jennifer J. Venditti

16 Dependency-based Phonologies 533
Harry van der Hulst

17 The Acquisition of Phonology 571
Katherine Demuth

18 Phonology as Computation 596
John Coleman

19 Using Psychological Realism to Advance Phonological Theory 631
Matthew Goldrick

20 Learning and Learnability in Phonology 661
Adam Albright and Bruce Hayes

21 Sign Language Phonology 691
Diane Brentari

22 Language Games 722
Bert Vaux

23 Loanword Adaptation: From Lessons Learned to Findings 751
Carole Paradis and Darlene LaCharité

References 779

Index 914

Contributors

Adam Albright, MIT

R. Harald Baayen, University of Alberta

Eric Baković, University of California – San Diego

Mary E. Beckman, Ohio State University

Diane Brentari, University of Chicago

Andries Coetzee, University of Michigan

John Coleman, Oxford University

Stuart Davis, Indiana University

Katherine Demuth, Macquarie University

Mirjam Ernestus, Max Planck Institute for Psycholinguistics

Matthew Goldrick, Northwestern University

John Goldsmith, University of Chicago

Matthew Gordon, University of California – Santa Barbara

Gunnar Ólafur Hansson, University of British Columbia

Bruce Hayes, University of California – Los Angeles

Larry M. Hyman, University of California – Berkeley

Sharon Inkelas, University of California – Berkeley

Darlene LaCharité, Université Laval

D. R. Ladd, University of Edinburgh

David Odden, Ohio State University

Carole Paradis, Université Laval

Joe Pater, University of Massachusetts – Amherst

Sharon Rose, University of California – San Diego

Elisabeth Selkirk, University of Massachusetts – Amherst

Harry van der Hulst, University of Connecticut

Bert Vaux, University of Cambridge

Jennifer J. Venditti

Rachel Walker, University of Southern California

Alan C. L. Yu, University of Chicago

Preface

This new volume on phonological theory is in some respects a continuation of the *Handbook of Phonological Theory* published by Blackwell in 1995. The present book was several years in the making, and reflects both the changes that the field has gone through in the years since the first handbook was written, and a shift in the precise character of the questions we hope to see answered in a book such as this. As you will see in the chapters that follow, we have asked each author to take a step back from the research that has been published over the last decade in each subfield in phonology, and to ask what the broader questions are that have been the focus of investigators over a longer period of time. Having identified the long-standing questions, the authors were then asked to pass judgment – as best they could – on the degree to which the field had succeeded in providing answers to these questions.

In this way, our handbook takes on a perspective that is different from many others in linguistics. We have asked our authors to set as their primary goal to provide some grounds for determining the degree to which phonology – as a whole, and as a set of subdisciplines – displays a cumulative character, which is to say, succeeds in asking questions that are both interesting and useful in some respects, and then – just as importantly! – answering them. In particular, we asked our authors to avoid as much as possible adopting the stance of the scholar who predicts where the field will, or should, go in the next five to ten years, and what the important open questions are. While there certainly is a place for such gazing into a well-focused crystal ball, we felt that the present handbook was not that place.

Comparing the present handbook to the one that was produced in 1995, we seem to find, too, that the field has expanded: it now includes a good deal more content and emphasis on phonetics, on variation, and on computational approaches. In reality, the growth is more a matter of perspective than anything else: studies on phonetics, variation, and computation that were of interest to phonologists have existed for a long time, but the perception is now much stronger that this work is not outside the field of phonology (though of interest to some phonologists),

as it is a real and integral part of the field itself. The broader range of the questions covered by the authors in this volume is testament to that change.

If we were to point to the greatest single difference between the work in the two volumes, it would have to be the considerable replacement of the analytic tools of phonological derivations within a generative framework with those optimality theories, utilizing ranked constraints from a universal inventory of violable phonological constraints. The first chapter in this volume, by David Odden, provides an illuminating overview of the nature of the questions which have been explored, with the goal of understanding the essential difference between these two approaches.

In Chapter 2, Eric Baković revisits the topic of opacity and examines its role in distinguishing ordered versus parallel phonological derivations. He demonstrates that the range of opaque relations between underlying and surface forms does not partition neatly into "counterfeeding" and "counterbleeding" classes and moreover that there are cases which fit Kiparsky's seminal definition of opacity that cannot be generated by ordered derivations.

Changes in analytic tools are often accompanied by shifts in perspectives. Age-old problems are given a fresh look while new puzzles come about, as novel theoretical tools are tested. In this respect, the emphasis on constraint interaction and monostratalism has certainly left an undeletable mark on how one thinks about the relationship between the morphological and phonological components of grammar. In Chapter 3, Sharon Inkelas surveys the pros and cons of a monostratal interpretation of the morphology-phonology interface, and details the many ways in which morphological processes can be sensitive to phonological information and vice versa, highlighting properties that any theory of the phonology-morphology interface must take into account.

In Chapter 4, Stuart Davis offers an overview of the development of moraic phonology and provides a survey of a wide range of linguistic phenomena where the mora plays an important role, including thorny issues such as the existence of moraic onsets and the replacement of moraic quantity with phonetically-defined weight sensitivity in language.

Matthew Gordon provides a broad overview of stress systems, including quantity insensitive systems and quantity sensitive systems, in Chapter 5. Gordon provides an account of what constitutes "weight" in various quantity-sensitive systems and discusses the relationship between word-level and phrase-level stress. The chapter also presents an in-depth comparison of foot-based and grid-based representations of stress and discusses their ramifications for models of stress.

John Goldsmith presents in Chapter 6 a synoptic overview of the ways of understanding the syllable that have played a role in phonological thinking over the last hundred years, emphasizing the ways – often complementary, and not always consistent – in which the different conceptions of the syllable have emerged and developed in discussions in the literature. The two most appealing approaches have based on waves of sonority, on the one hand, and constituent structure as developed by mid-century syntacticians, on the other. A few phonological theories

have tried to jettison the syllable, but rarely with any lasting success, and a successful synthesis of the best of what has been learned still awaits us.

Larry M. Hyman has studied tone languages – mostly, but not exclusively, African tone languages – in great depth since the 1960s, and in Chapter 7, he offers the reader a rich account of many of the properties of tone languages that have emerged in studies over the past several decades. He asks what we have learned about how tone is different from other aspects of spoken language, and how it can nonetheless shed a great deal of light on the way in which phonological information is organized in natural language.

Sharon Rose and Rachel Walker provide a thorough overview of harmony systems that includes vowel harmony, consonant harmony, and vowel-consonant harmony in Chapter 8. They provide an account of the triggers and targets of harmony in the case of continuous sequences and when harmony acts at a distance. For the latter, they provide an analysis of segments that block harmony when they intervene between trigger and target and those that are transparent to harmony. They identify a broad dichotomy between consonant harmony on one hand and vowel harmony – including vowel-consonant harmony – on the other that is framed in terms of blocking and transparent segments and the functional grounding that provides insight into why consonant harmony does not, in general, admit transparency, while harmony with vowels does. Finally, they discuss a range of fundamental issues, including the domain of harmony, directionality, and locality.

The notion of contrast reduction has been central to many major developments in phonological theories. Yu's chapter, which is an expanded version of an article titled "Mergers and Neutralization," that appeared in the *Companion to Phonology* (Wiley-Blackwell 2011), provides an overview of the range of contrast reduction phenomena in the world's languages and past theories that try to explain the typological tendencies. Yu places a particular focus on the problems raised by covert contrasts (i.e. incomplete neutralization and near mergers). He questions the reliability of the traditional methods of phonological investigation (see also Ladd's chapter) and argues for the need to evaluate the presence and absence of a phonological contrast at a more nuanced level.

While it is undeniable that languages are products of history, the issue of how phonological explanation should take into account historical factors remains a contentious one. Hansson's chapter, which originally appeared in the journal *Language and Linguistics Compass*, reviews an wide array of theoretical stances that phonologists have taken over the years, ranging from strictly modular approaches to the more integrationalist. Hansson shows that this controversy largely stems from questions about the nature of sound change and what models of sound change reveal about the nature of phonological knowledge.

D. Robert Ladd's chapter on the role of phonetics in phonology is a good example of how the thematic questions at the center of phonological discussions have evolved over the last 15 years. The time-depth of his discussion, involving scholars working over almost all of the twentieth century, is considerably deeper than that found in any of the chapters in the 1995 volume, and Ladd explicitly

draws together the views that Trubetzkoy developed in the 1930s with those at the heart of classical generative phonology and those that scholars today are developing, often under the influence of far richer computational resources than was imaginable even 25 years ago. The easy assumptions that phonetic reality can be modeled with a well-designed symbolic representation, such as that produced by the International Phonetic Association, have been widely challenged, and Ladd asks what alternative empirical accounts are available to us now for characterizing the nature of phonetic reality.

As noted earlier, a major change since the last edition of this handbook has been the rise in prominence of phonetic, variationist, and computational approaches in phonological investigation. A clear reflection of this is in the greater willingness on the part of many phonologists to engage data sources that have not played a large role in early theoretical developments. In their contribution, Ernestus and Baayen review findings of recent corpus-based studies of sound patterns and highlight important lessons to be learned from such studies. The appearance of what might in former times be thought of as "messy data" in the phonological discourse has invited renewed discussion on the abstractness of phonological knowledge, which the authors integrate by comparing the merits of abstraction-based vs. exemplar-based models of phonology.

In Chapter 13, Andries Coetzee and Joe Pater discuss several theoretical approaches to variation in phonology. It is fair to say that the emergence of widespread interest in variation among theoretical phonologists is one of the more significant changes in the field at large since the publication of the 1995 *Handbook of Phonology* (which did not contain a chapter on variation). Coetzee and Pater review a range of proposals in which variation is taken to illuminate the core phonological grammar rather than obscure it. Instead of regarding variation as a performance-related epiphenomenon that must be factored out in order to characterize the phonological grammar, they focus on understanding the locus (or loci) of variation in the grammar and the empirical consequences of various assumptions in this regard.

Lisa Selkirk has been doing influential work on the interface between phonology and other components of the grammar for over three decades. In Chapter 14, she discusses the interface between phonology and syntax in terms of the relationship between syntactic constituents and prosodic constituents. She presents a thorough account of the way that prosodic constituent domains for phonological and phonetic phenomena at the sentence level are related to syntactic constituency.

A domain where interests in phonetic and phonological investigations have converged in recent years is the area of intonational research. In their chapter on intonation, which is a slightly revised version of their contribution in the *Handbook of Phonetic Sciences* (edited by Hardcastle, Laver, and Gibbon, 2009), Beckman and Venditti review the development and advances of experimental intonational research and highlight their contributions to the understanding of intonational phonology and prosodic typology.

In Chapter 16, Harry van der Hulst presents an overview of work on government- and dependency-based phonology, which explores the consequences for phonological

theory of developing phonological representations that incorporate in an essential way formal asymmetrical relationships between abstract elements. "Asymmetrical" here refers to the important differences between what are called the head and the dependent, connected by a relation of dependency. Van der Hulst reviews recent work in this area, and notes respects in which government phonology has brought out parallels involving relations between elements in syntax and in phonology.

Katherine Demuth discusses in Chapter 17 the ways in which contemporary phonological theory has been reflected in the research concerns of a large part of the language acquisition community. Among the themes whose importance has grown over the last two decades are the relevance of surface-oriented phonological patterns, of prosodic patterns at both the syllable and foot level, of markedness and underspecification, and frequency. At the same time, conclusions can be drawn that are more robust in light of the wider range of languages that have been studied by acquisition researchers.

In Chapter 18, John Coleman guides the reader through the developments over the last 50 years which have influenced phonological modeling, bringing out the often only tacit connections between computational conceptions and phonological analyses, such as those employing finite-state methods, purely declarative formalisms, or techniques based on neural networks.

As phonologists rely more and more heavily on experimental methodologies, the question of the psychological status of phonological constructs becomes ever more important in the analyst's mind. Goldrick (Chapter 19) explores the notion of psychological realism in phonological inquiries, highlighting the need to differentiate at least three levels of analysis: functional, algorithmic, and neural. Using well-formedness judgments as a case study, he emphasizes the need to articulate in greater specificity the functional architecture of language processing in the context of interpreting experimental results.

Adam Albright and Bruce Hayes present an account of phonological learning in Chapter 20. They focus on formal systems designed to model the path by which children acquire the phonological grammar of their first language and evaluate the adequacy of the systems in terms of their ability to elucidate what is known about linguistic competence in three specific areas. The areas that they give special attention to are phonotactic knowledge, phonological alternations, and patterns of variation. They argue that any system capable of mimicking human performance in these areas – including the mistakes – will have reverse-engineered key aspects of the phonology of natural language in a way that enriches our understanding of both theoretical phonology and the broad character of observed phonological phenomena.

Diane Brentari's chapter on the phonological structure of sign languages extends her chapter on the subject in the 1995 edition of the handbook. In the current chapter, she reviews our better understanding of three important aspects of sign languages: their phonological structure, their iconicity (that is, the principles and patterns relating phonological structure to real or understood world structure), and the respects in which the phonologies of sign language are influenced by the

physical modality used, notably the structure of the signing hands and body, and vision, which is the perceptual system used for perception.

The last two chapters focus on linguistic evidence that has often been taken to be extralinguistic. Both sets of authors instead argue for the centrality of such evidence in testing and advancing phonological theories. In his chapter on language games (Chapter 22), Vaux provides an overview of the empirical and theoretical advances language games have contributed to phonological research. He argues that research on language games not only reveals subtleties of phonological representation, they also shed light on the cognitive limits of linguistic operations and language acquisition, as well as architectural issues such as opacity. In their chapter on loanword adaptation, the last chapter of this book, two veteran loanword phonologists offer a summary of major findings in loanword adaptation research, and reflect on major lessons learned from this line of inquiry. Echoing Ernestus and Baayen's call for more corpus studies in Chapter 12, Paradis and LaCharité illustrate in Chapter 23 the importance of corpus construction and the need to pay attention to statistical generalizations with their own Project CoPho loanword database.

We offer these chapters to both the reader who is relatively new to the field and to the expert knowing full well that no-one can keep fully up to date on all the fields that now comprise phonology. We thank our authors for their efforts, as well as for their patience and forbearance during the book's preparation, and we hope that our readers will profit from the chapters as much as we, the editors, have.

We would like to dedicate this book to the memory of G. N. Clements, who was planning to contribute a chapter to this handbook, and who left us too soon to able to do so. Like many others, we admired Nick's work and were influenced greatly by that work, and we will miss him.

1 Rules v. Constraints

DAVID ODDEN

1 Background

The goal of a theory of phonology is to elucidate the nature of "phonology" at a conceptual and predictive level. The title of this chapter refers to a comparative evaluation of rules and constraints as successful theories of phonology, which implies having a standard of evaluation, and adequate clarity as to what "rules" and "constraints" refer to. Neither prerequisite is trivial to satisfy.

1.1 The Scope of Inquiry

Certain assumptions about the nature of phonology must be considered, even lacking agreement on which assumptions to make. First and foremost, deciding whether phonology is based on rules or constraints, or a mix of the two, requires having objectively expressible statements of phonologies within different frameworks whose consequences can be compared. Therefore the theories must have a definite form, that is, they must be formalized. The entities which make up a phonological grammar should be expressions, which are finite sequences of elements taken from a specified set, and combined by rules of construction that define well-formed statements of rule or constraint. The value of formalism is its power to make objectively-interpretable statements about the phonology which can be checked against fact. To evaluate rules versus constraints as models, we should then consult the formalisms of the theories, to see whether one theory better passes the test of empirical and aesthetic adequacy.[1] Problems in this area are not trivial; certain theories of constraints or rules are severely under-formalized so that it is hard to know what predictions the theory makes; and a number of

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

theories are under-applied in the sense that it is impossible to determine from examples how particular phenomena would be analyzed.

Assuming that we are comparing formal theories, we must resolve questions about the scope of phonology, including how much of "phonetics" or "morphology" is phonology, and whether all facts bearing on phonology are the responsibility of the theory. Generative phonology traditionally encompasses a broad range of processes which might be considered phonetic (allophonic) or morphological (rules with lexical or morphological conditions), but the edges of phonology may also be contracted for theoretical purposes, viz. restrictiveness. Thus Webb (1974: 127) excludes metathesis from phonology, stating that "synchronic metathesis is not a phonological process. In the residual cases of metathesis, the rule is always morphologically restricted," enabling the "Weak Metathesis Condition," a restriction against reordering in phonology. If phonology is deemed to be concerned only with biuniquely recoverable surface-true relations between sounds (e.g. allophonic vowel nasalization in English), and abstract phonological alternations are to be described by the formal methods of morphology, a theory designed to account for just surface phonotactics cannot be meaningfully compared to one designed to account for both phonotactics and abstract morphophonemics.[2] A surface-phonotactic view of phonology thus must ignore a substantial portion of research into phonological grammars, on Bedouin Arabic (Al Mozainy 1981), Finnish (McCawley 1963; Harms 1964; Karttunen 1970; Keyser and Kiparsky 1984; Kiparsky 2003a), Chukchi (Krauss 1981), Kimatuumbi (Odden 1995), Klamath (Kisseberth 1973; White 1973), and Ojibwa (Piggott 1980), and numerous other languages.

There are also questions as to the level of explanation demanded of a theory – do we demand formal explanation, or formal and functional explanation? Much of the course of phonological theorizing has involved the increasing absorption of substantive factors into the theory, in an attempt to narrow the gap between prediction and observation. Comparative evaluation of theories implies determining which theory is better at making definite the notion "possible rule" or "possible constraint." The notion "possible" is used in two ways. One sense is theoretical well-formedness, that is, a rule constructible by free combination of elements, according to a theory of the form of rules. In that sense, "A→B/C__D" would be a possible rule, but "→B__/ACD" would not. McCawley (1973: 53) points to a different sense, the metaphysically possible, claiming "One who takes 'excessive power' arguments seriously has as his goal characterizing 'phonological rule' so as to include all and only the phonological rules that the phenomena of a natural language could demand. . . ." This notion of "possible rule" seems to mean what does exist, so *is* attested, or that which we have solid scientific or philosophical reason to conclude *must exist* now or in the past or future, just waiting to be discovered. The latter kind of "possible" depends on metatheoretical expectations, so McCawley intuits that assimilation of nasal to labials alone is not a possible rule (the present author does believe that such a rule is possible, if unlikely).

Whether such a rule is possible is not central to this discussion: what is essential, is distinguishing the undiscovered from that which is impossible by the nature

of language. Expansion of the substantive content of phonological theory narrows the predictive gap, though, complicates the theory and renders it redundant with respect to the extragrammatical physical explanations for the gap. If phonology is only a system of symbolic computations where the syntax of computations defines a broad class of possible rules, and separate aspects of languages referring to substance (perception, acoustics, articulation, language learning, and the transduction between grammar and linguistic behavior) explain why some formally allowed rules have negligible probability of attestation (as argued by Hale and Reiss 2008; Morén 2007), then failure to capture a generalization about substance within the theory of computation is not an argument against the theory of computation. But there is no universal agreement that the object of investigation *is* the computational apparatus rather than the full and undifferentiated panoply of factors influencing linguistic sound.

A second metatheoretical question affecting a comparison is whether phonology describes abstract string collections, or the mental faculty which generates them. If phonology only models strings, then considerations such as the results of psycholinguistic tests or problems regarding infinities in the model – infinite sets of candidate or sub-rules – are irrelevant to theory selection.[3] An example of how different conclusions are reached depending on whether one considers just the strings, versus the strings plus the mechanisms, is Mohanan (2000: 145–146) versus Calabrese (2005: 34). Mohanan contends that a rule [+nasal] → [+voice] is "logically equivalent" to a negative constraint *[+nasal,–voice], while Calabrese contends that rules and constraints are totally different means of implementing a linguistic action and are ontologically different. Mohanan is correct that the rule and the constraint describe the same string classes – are weakly equivalent; Calabrese is right that the imputed mental mechanisms of rules versus constraints are different – are not strongly equivalent.[4]

Even if we presume that phonology should be concerned with a mental faculty as well as the sets of strings, we must also determine whether phonology is concerned with all sound-related behavior, or just that behavior which generates the strings. A mentalist view of phonological grammars would care whether insertion of [i] after a word-final obstruent is regulated by a rule or a constraint, and whether this takes place in a single step or many steps; but a mentalist view of phonological grammars does not automatically care about the behavior of speakers of such a language under certain types of psycholinguistic testing, since a mentalist view of grammar does not automatically hold that all aspects of the mind pertaining to language sound are contained in the phonological component of a grammar.

To properly contrast "rules" versus "constraints" in phonology, we must also determine what these terms refer to, because we want our conclusions about differences between rules and constraints to reflect the concepts themselves, and not quirks of particular theories of rules or constraints. Many definitions of "rule" are offered in the Oxford English Dictionary, but the ones that seem closest to its linguistic use are:

A fact (or the statement of one) which holds generally good; that which is normally the case.

A principle regulating the procedure or method necessary to be observed in the pursuit or study of some art or science.
(Grammar). A principle regulating or determining the form or position of words in a sentence. In modern linguistics, usually applied to any one of a system of rules that can be formulated in such a way that together they describe all the features of a language.

The closest applicable definition of constraint is

The exercise of force to determine or confine action; coercion, compulsion.

In addition, the terms "principle," "condition," and "convention" are often used in linguistics to describe what often seems to be the same thing as a constraint, perhaps with the implication of greater generality or a stronger commitment to universality.

In other words, the terms "rule" and "constraint" have developed into terms of art in linguistics, requiring special definition, and the ordinary meanings of the words only have an approximate correspondence to their linguistic use. The original formal notion of a "rule" derives from the computational notion of Post production systems, developed in the 1930s by Emil Post (Post 1943). In generative grammatical theory, the essential characteristic of a *rule* is that it maps classes of strings onto other classes of strings in a specific way: the rule encodes the particular change. Classically, rules in generative grammar also have the Markov property, that the device or rule refers only to its current state (the input string) and not some future or past state or string – such a device is "Markovian."[5] Thus a rule which states "AXB → AZB" means "if you find a string analyzable as AXB (at the current stage of the derivation), it maps to AZB (at the successor stage)." A non-Markovian rule could refer to facts of a prior stage in a derivation.

Constraints are less well-defined largely due to the fact that their primary characteristic is "not being a rule."[6] A constraint is essentially a "limit," so the exact nature of a constraint depends on whether one is constraining a rule, a derivation, or a representation. Contemporary usage sees constraints as evaluating structures, but originally, constraints were limits on rules, typically defined in terms of a string property. The property of "overarching, non-local influence," that is, relevance to something more than one rule, is another behavioral characteristic of constraints. Constraints can be either Markovian (morpheme-structure or surface well-formedness constraints, which state generalizations at one level) or non-Markovian (transderivational constraints on input-output relations, OT Correspondence Constraints, the Elsewhere Condition), but are typically not seen as *holding of* just one rule or step in a derivational mapping, assuming derivations.[7] The general concept "constraint" does not say whether the mechanics of grammar allow constraints to be violated, and says nothing about how constraints are enforced or how potential or actual violations are handled. Constraint-based theories differ considerably in this respect, some theories (Declarative Phonology) disallowing violations of constraints, others (famously, OT) allowing them.

To anticipate the results of this investigation, there is no substantial difference between rules and constraints per se in their power to deal with phonological systems. The important differences reside in properties of particular theories of rules and constraints. Different theories of rules and constraints combine simple theoretical properties in many ways. For example, "surface-trueness" is a property sometimes associated with constraints and not rules, but some rule theories require the rules of language to be surface true (Natural Generative Phonology; Equational Grammar, Sanders 1972b), and OT is founded on the idea that constraints can be violated. The most important properties of the formal statements used in rule or constraint systems which we will be watching for are:

Globality: the statement applies "generally" in a language, not just at one point.
(*Language*) *Universality*: the statement pre-exists in UG: is not dependent on exposure to a particular language.
Inviolability: the statement must be true of particular levels of representation.
Negativity: the statement may give conditions that must not hold.
Ordering: the statement interacts with other statements according to language-specific priority.
Multiple Representations: more than one representational string is involved in computing the output form.

1.2 The Seeds of the Rule/Constraint Distinction

While the idea of directly and literally stating all of the facts of the mapping performed by a rule within the formalization of the rule itself would seem to characterize rule-based grammar, such a theory has never existed, and generative grammar has always operated with local rules and global meta-principles of rule interpretation. Nevertheless, the development of the concept "rule" in generative grammar from the most direct and literal statement of string-to-string mapping inevitably gave rise to the separate concept "constraint," when linguists faced recurring linguistic regularities which were not easily expressed in a general-purpose symbol-manipulation algebra. In saying that rules map classes of strings onto classes of strings, we recognize that rules use abbreviatory expressions to reduce classes of objects to compact symbols, for example a symbol to represent "consonant" or "NP." Rules are not written to apply exclusively to particular concretes such as [f] or *the child*. Formal linguistic statements are necessarily written with an abbreviatory notation referring to linguistic objects, and conventions that transcend a specific rule must be established for interpreting rules.

The development of the distinction between rules and constraints began in syntax, and early concepts of phonological constraints were a direct consequence of the prior development of such ideas in syntax – the implicit goal is to develop a theory of grammar. Early generative grammar as exemplified by Chomsky (1957, 1965) depended heavily on rules which explicitly stated the operations performed. Thus the Particle Shift transformation in Chomsky (1957: 112) is stated as "X-V1-Prt-Pronoun → X-V1-Pronoun-Prt," that is, when a particle precedes a pronoun,

the pronoun obligatorily moves to precede the particle: a separate optional rule addresses the situation where the word after the particle is a full non-pronominal NP. In this rule, X is taken by general mathematical convention to be a variable representing "any sub-string." Chomsky considers (p. 76) but does not formalize a generalization to the effect that ordinarily optional Particle Shift is obligatory *if* the post-verbal nominal is a pronoun, setting the stage for higher-order "conditions" on rule application separate from classical string-rewriting rules. Such a generalized version with an "obligatory if pronoun" condition does not follow the simple string-rewriting model, indicating that something in addition to string-rewriting statements are required.

A principle of Chomsky (1964: 931), dubbed in Ross (1967) the "A-over-A principle," gave rise to the first explicit constraints in generative grammar. This principle asserts that "if the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW)." That is, when category A dominates an A, how is reference to "A" in a rule interpreted with respect to a string – as applying to the higher A or the lower A? According to this principle, interpretation of "A" is limited to just the higher A. A-over-A is not a rule (it does not state a string mapping), and it is global rather than local. It thus had a separate status, as a limitation on grammars, and an autonomous and universal claim about the notion "rule of grammar."

The consideration of factoring generalizations out of rules and giving them independent status – the globality property – took on a major role in linguistics with Ross (1967), who argues for the unambiguous necessity of autonomous constraints in grammar, in order to account for the facts covered by A-over-A. Ross argues that greater generality and simplicity can be achieved by removing certain considerations from explicit rule statements, and giving them the status of separate limitations or constraints on grammars. Since a rule is one derivational mapping, the only means of propagating a formal identity across rules in early generative rule theory was via a convention which defines a notation, for example, "X *means* a string of symbols of unbounded length." Ross-constraints change the conception of language because those statements cannot be reasonably construed as "defining the meaning of formal symbols," but they also are not linearly ordered string-rewrite rules.

The first constraint postulated by Ross is S-pruning (p. 26): "delete any embedded node S which does not branch . . . ," motivated by the fact that syntactic theory at that time held, counter-intuitively, that "his" and "yellow" in "his yellow cat" are sentences. Ross comments (emphasis added) "This principle should *not be thought of as a rule* which is stated as one of the ordered rules of any grammar, but *rather as a condition upon the well-formedness* of trees, which is *stated once in linguistic theory*, and applies to delete any non-branching S nodes which occur in any derivations of sentences in any language." In terms of globality and the statement of well-formedness, S-pruning has clear affinities to a constraint, but insofar as it also includes a statement of repair – the principle is not interpreted to mean "block a rule that *would* create such a structure" – S-pruning resembles

a rule. Other constraints such as the Complex NP Constraint – "No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation" – exert a blocking influence, preventing wh-movement from generating *"Who does Phineas know a girl who is jealous of."

The constraint-based tactic, best summarized in Ross (1967: 271), is "that many conditions previously thought to be best stated as restrictions on particular rules should instead be regarded as static output conditions, with the rules in question being freed of all restrictions": recurring aspects of multiple rules can be factored out and stated separately, making the formal statements of the rules simpler. Extraposition from NP thus need not explicitly list the content of its right-edge variable, to block the sentence *"That a gun_i went off surprised no one which I had cleaned $__{}_i$." Instead, this effect is achieved via a rule-independent principle – a constraint – on the content of variables in certain kinds of rules. Constraints might be universal (the Coordinate Structure Constraint was claimed to be universal) or language specific (the Pied-piping constraint is language specific).

Constraints typically had two realizations in early generative grammar, blocking and filtering. The blocking function says that if a particular rule application would contradict some constraint, the rule could not apply. Ross's Coordinate Structure Constraint thus blocks wh-movement from applying to "Bill and who bought biscuits?" The notion of "filtering" is brought out explicitly in Chomsky (1965: 137–139), to explain why the relative clause and higher NP must contain identical nouns, to prevent an unrealizable deep structure [the man [Bill saw the woman]]. Chomsky notes (pp. 138–139) "The transformational rules act as a 'filter' that permits only certain generalized phrase-markers to qualify as deep structures." Blocking and filtering are not particularly distinct when applied to optional rules (as syntactic rules have sometimes been held to be), and blocking an optional rule is string-equivalent to freely applying the rule and then filtering out violations of the constraint. Constraints and filtering achieved greater prominence in such works as Ross (1967), Emonds (1970), Perlmutter (1971), Hankamer (1973), Lakoff (1973), for instance and, as we will see below, a number of works in phonology.

2 Rules in Phonology

The concept of a (synchronic) generative phonological rule was developed in such works as Chomsky (1951), Halle (1959b, 1962, 1964), Chomsky and Halle (1965), Kiparsky (1965), Lightner (1965), McCawley (1965), Schane (1965), Zwicky (1965), Sloat (1966), Harris (1967), culminating in the essential reference work in the theory of generative rules in that era, Chomsky and Halle (1968). In this theory, often called the SPE ("Sound Pattern of English") theory, a grammar is a linearly ordered sequence of rewrite rules mapping an underlying form (the output of the syntax) to the surface representation.

The main theoretical concerns of phonology were the sub-theories of ordering, features, and rule formalism. All three aspects must be considered in evaluating

the theory against its competitors. Representation and rule statement are closely related since rules map between representations. Ordering bears on the question since some constraint-based theories preclude ordered derivational steps, and because a rule implies at least two levels, the input and output.

2.1 Rules and Conventions

A grammar is a linearly ordered sequence of rules, and, as is characteristic of generative formalism at the time, a rule is defined (Chomsky and Halle 1968: 391) as:

> *Z X A Y W* → *Z X B Y W*, where *A* and *B* may be φ or any unit; *A* ≠ *B*; *X* and *Y* may be matrices; *Z* or *W* may be C_i^{∞} for some *i*; *Z*, *X*, *Y*, *W* may be null; and where these are the only possibilities.

Feature matrices identify sets of segments by conjoining specified features, thus the expression [+high,–voice] refers to the set of all segments which are both *+high* and *–voice*. Since the vast majority of phonological rules operate on just a single segment at a time, rules were usually stated in a format that factors out the non-changing segments, thus B → C / X___Y where X, Y could be any string of matrices, and B and C are a matrix or the null string.[9]

Given this characterization of rule, any mapping from specific string to specific string is possible (meaning, allowed by the syntax of rule construction) – a rule *mowzɪz* → *mɪdl̩tawn* is a possible rule, and so is the following, which refers to classes of string:

(1) X [+syllabic] [+nasal] Y → X [+syllabic,+nasal] [+nasal] Y

However, not every mapping of string *class* to string *class* is possible. Feature theory defines possible matrices, and given the nature of SPE's feature theory, the set {æ, m, š, ʊ, g} cannot be referenced to the exclusion of {a, n, i, ɪ, u, s, b, p, t, k}, so no rule can effect the mapping:[10]

(2) $\{æ, m, š, ʊ, g\}_i$ → $\{š, ʊ, g, æ, m\}_i$ / {a, n, i, ɪ, u, s, b, p, t, k} ___

That is, even though any rule (as defined above) is possible, not every imaginable mapping of string class to string class is a possible rule in the theory. A rule in SPE is local (not global), not universal, positive (not negative); rules are linearly ordered, there can be multiple representations (a derivation), and while rules are not violated in the immediate output of the rule (modulo lexical exceptionality and optionality), they need not be true of any level.

The notion of "rule" becomes more complex because in SPE, sets of elementary rules can be combined into rule schemata via auxiliary expressions, for the purpose of grammar-evaluation and ordering. The notion of "evaluation" plays a significant role in grammatical theory – the assumption is that children learning a language are faced with multiple competing hypotheses which need to be

evaluated, the best one being the one actually acquired. The claim of the theory is that when rules resemble each other in specific ways, this resemblance is a significant linguistic generalization which needs to be captured. For example, a grammar could contain the following pair of elementary rules:

(3) $[+A] \rightarrow [-D] \; / \; __ \begin{bmatrix} +E \\ -F \end{bmatrix}$

$[+A] \rightarrow [-D] \; / \; __ \; [-G]$

The similarity between these rules can be captured via a notational device, the brace notation, whereby a single statement can express these two elementary rules:

(4) $[+A] \rightarrow [-D] \; / \; ____ \left\{ \begin{bmatrix} +E \\ -F \end{bmatrix} \right\}$
$[-G]$

which means "Any segment which is [+A] becomes [–D], when it stands before either a [+E,–F] segment or a [–G] segment." The significance of such abbreviation is two-fold. First, the evaluation metric assigns a greater value to a sequence of rules which can be collapsed via an abbreviatory convention than a similar uncollapsible rule sequence, and second, sub-rules abbreviated with abbreviatory devices apply disjunctively,[11] so only one of the rules in a schema can apply to a given segment. The evaluative function of abbreviatory notations was the most important, because language acquisition was seen as the process of selecting the formally simplest grammar consistent with the data. Abbreviatory devices then say that certain sets of rules are simpler in the sense that their "cost" is a fraction of the cost of the total set of individual rules. The mappings described as {æ, m, š, ʊ, g} → {š, ʊ, g, æ, m} / {a, n, i, ɪ, u, s, b, p, t, k} ___ can only be accomplished via a highly disvalued list of unreducible changes æ → š / a__ ; æ → š / n__ ; m → ʊ / a__ ; etc.

Other devices were employed to express optional elements, so the context "___([+A])[–C]" means "when the segment precedes something that is [–C], with or without one intervening [+A] segment," and "___$[+A]_0$[–C]" means "before a [–C], with any number of intervening [+A] segments." Another significant device was the feature-coefficient variable, typically expressed with Greek letters α,β,γ . . . which represented the two feature values {–,+}. This notation was widely used to express assimilation processes, such as the following place assimilation for nasals.

(5) $[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha\text{ant} \\ \beta\text{cor} \end{bmatrix} \; / \; ___ \begin{bmatrix} -\text{syl} \\ \alpha\text{ant} \\ \beta\text{cor} \end{bmatrix}$

This abbreviates the following four rules.

(6) a. [+nasal] → $\begin{bmatrix}+\text{ant}\\+\text{cor}\end{bmatrix}$ / ___ $\begin{bmatrix}-\text{syl}\\+\text{ant}\\+\text{cor}\end{bmatrix}$ c. [+nasal] → $\begin{bmatrix}-\text{ant}\\+\text{cor}\end{bmatrix}$ / ___ $\begin{bmatrix}-\text{syl}\\-\text{ant}\\+\text{cor}\end{bmatrix}$

b. [+nasal] → $\begin{bmatrix}+\text{ant}\\-\text{cor}\end{bmatrix}$ / ___ $\begin{bmatrix}-\text{syl}\\+\text{ant}\\-\text{cor}\end{bmatrix}$ d. [+nasal] → $\begin{bmatrix}-\text{ant}\\-\text{cor}\end{bmatrix}$ / ___ $\begin{bmatrix}-\text{syl}\\-\text{ant}\\-\text{cor}\end{bmatrix}$

Various aspects of the theory of rule formalism and schemata are set forth in SPE, especially pp. 393–399 for rule schemata, including X_0, X* and other notations. See also Bach (1968) for the Neighborhood Convention notation.

The complement notation suggested in Zwicky (1970) introduces "negativity" into rule statements which otherwise state what must hold for a rule to apply, since the complement notation refers to "anything but," that is, what must not hold, for a rule to apply. An example of that kind is the *ruki* rule of Sanskrit, where /s/ becomes [ṣ] after the class *r,u,k,i*, provided that the following segment is not /r/. The right-hand context could be expressed "–[+son,–nas,+cor]" or "~[+son,–nas,+cor]" with the complement notation. As Zwicky notes, the complement of a natural class – a feature conjunction – is, by DeMorgan's law for negation of a conjunction, equivalent to a disjunction of negated values (¬(A∧B)≡(¬A∨¬B)), thus the right-hand condition can be stated as {–son,+nas, –cor}. A simple translation between direct statement of context and complement statement is possible for a single matrix being a blocking context, but not for a segmental sequence. Suppose a rule applies after certain segments but is blocked when immediately followed by [ba]. Simply changing conjunction to disjunction and reversing signs on the right-hand context does not give the desired effect. Such a conversion applied to the expression:

(7) $-\begin{bmatrix}+\text{voice}\\-\text{cont}\\-\text{nas}\\+\text{ant}\\-\text{cor}\end{bmatrix}\begin{bmatrix}+\text{syl}\\+\text{low}\end{bmatrix}$

would give:

(8) $\begin{Bmatrix}-\text{voice}\\+\text{cont}\\+\text{nas}\\-\text{ant}\\+\text{cor}\end{Bmatrix}\begin{Bmatrix}-\text{syl}\\-\text{low}\end{Bmatrix}$

which means "anything besides [b] followed by anything besides [a]." The difference in the two expressions lies in the fact that with the complement notation, the sequence [bi], [da] on the right would not block the rule, but with the negated

disjunction approach, such sequences would block. This points to an important question about blocking conditions, namely, does blocking ever require the characterization of a sequence of segments, or do blocking effects always involve the complement of a single element? A further point about blocking effects is that the negated disjunction statement presupposes the brace notation, and the validity of the brace notatation in phonology has been called into question, for example, by McCawley (1973). The connection with constraints should be clear, since a rule that applies except when a configuration is present is extensionally equivalent to one subject to an output condition, that is, a constraint against the configuration blocks the rule.

The SPE-era abbreviatory conventions were received skeptically: see McCawley (1973) for discussion. An important question raised there is whether the notations do, as claimed in SPE, represent sets of independently-existing sub-rules – the various sub-rules actually exist in the grammar and are simply evaluated as a single unit – or are the notations first-order concepts? The notations which abbreviate infinite set (X* and X_0) cannot represent the collapsing of sets of rules in a grammar at least under a "model of the mind" view of grammar since a mental grammar cannot contain an infinity, so some of the SPE notational conventions must be primitive and not abbreviatory.

McCawley proposes, regarding feature variables, that the notion of feature identity should be a first-order concept in rule theory, so that a rule assimilating the coronality value of segment 1 to that of segment 2 would encode this as "coronal(1) → coronal(2)," meaning "the value of coronal for 1 becomes whatever it is for 2." The significance of this change to the theory is that it narrows the gap between observation and formal prediction, ruling out a large class of rules which are expressible in the SPE notation, such as:

$$(9)\quad [+\text{syl}] \rightarrow \begin{bmatrix} \alpha\text{hi} \\ \beta\text{low} \\ \gamma\text{back} \\ \delta\text{round} \end{bmatrix} / ___ \begin{bmatrix} +\text{syl} \\ \delta\text{hi} \\ \alpha\text{low} \\ \beta\text{back} \\ \gamma\text{round} \end{bmatrix}$$

where features and values are mismatched. See Reiss (2003) and Section 2.4 for further discussion.

The main objection to the abbreviatory devices proposed in SPE is that large classes of non-generalizations could be expressed. The "dash-factoring" notation (p. 338):

$$(10)\quad X \rightarrow Y / \left[\frac{\quad}{Z} \right] Q$$

which means "Before Q, anything that is X becomes Y when it is also Z" was also little-used, and was seen as a spurious economization, being extensionally equivalent to the expression "anything that is both X and Y." Apart from being

a capricious "use it if you want" device, this device was used to coerce collapsibility in rules that could not otherwise be formally collapsed, such as the SPE Tensing rules (Chomsky and Halle: 241).

The star-parenthesis notation was motivated in that it was used to express a fact of language, but was supplanted by the theory of rule iteration (Howard 1972; Jensen and Stong-Jensen 1973, Kenstowicz and Kisseberth 1977). Angled brackets were employed for various purposes, primarily structure-preserving side-effects (e.g. in Slavic velar palatalizations where *k* becomes [č] but /g/ becomes [ž] and not [ǰ]). The brace notation was also viewed with skepticism, especially since the majority of recurring uses pertained to syllable structure and typically involved finding a way to make {C,#} be a natural class. The parenthesis, subscript-zero and variable feature notations were fairly well motivated in that phenomena which the devices were predominantly used for are not easily deniable. These notations still posed significant predictive problems. For example, factoring a string into units of two for stress purposes was not difficult (see (11a)) and appropriately so because binary stress units are well attested, but it was no harder to factor strings into groups of seven, thus the formal theory overgenerates.

(11) a. $V \rightarrow$ [+stress] /# C_0 $((VC_0)^2_2)_0$ ___
b. $V \rightarrow$ [+stress] /# C_0 $((VC_0)^7_7)_0$ ___

Nasal place assimilation (5) is evaluated the same as the unattested rule (12).

(12) $[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha\text{ant} \\ \beta\text{cor} \end{bmatrix} / ___ \begin{bmatrix} -\text{syl} \\ \beta\text{ant} \\ \alpha\text{cor} \end{bmatrix}$

The class of attested rules of natural languages that motivate feature-variable notation seems to be a small fraction of the set of predicted rules, which is quite problematic if the theory is held responsible for distinguishing "actual languages" from "non-languages." The advent of nonlinear phonology seemed to eliminate the motivation and need for these notations (though see below), where a different theory of representations resulted in the possibility of expressing the facts at least as well. A similar trade-off between representational richness and statement-impoverishment is to be found in certain constraint-only theories, including Candidate Chains in OT and Declarative Phonology.

2.2 Blocking and Repairing Conventions

While the SPE theory with abbreviatory notations does a remarkable job, by comparison to previous formal theories of phonology, in characterizing possible versus impossible grammars and matching that to attested languages the theory mispredicted the possibility or probability of phenomena. Some of this stems from the substance-free nature of formalism, which counter-intuitively puts palatalization before back vowels and palatalization before front vowels on an equal footing.

On the assumption that this should be addressed by the formal theory, SPE introduced a major departure from strict rule theory, via a set of universal "rules" (not part of a grammar: p. 403), namely the markedness rules which encode aspects of phonetic substance. Given the device of linking, these rules automatically and globally modify the immediate output of rules. This introduces the notions of automatic repair and persistent rule, which played a major rule in the operation of non-linear phonology.

Under the markedness and linking proposal, lexical representations may have the values "u" (unmarked) or "m" (marked), which map to plus and minus by universal rules such as [*u*low] → [–low], [*u*del.rel] → [+del.rel]/[__–ant,+cor] (pp. 419–435). These rules also apply to the output of phonological rules, so given a rule changing F, a feature G whose unmarked value depends on F may be reassigned by a markedness rule. In Slavic, the rule [–ant] → [–back] / ___ [–cons, –back] derives /k g x/ → [č ǰ š]. Without markedness rules, this would only result in *[k^{y} g^{y} x^{y}]. A direct statement of the actual change requires more complex formulation with angled brackets (which encode discontinuous dependency not expressible via parentheses):

$$(13)\quad \begin{bmatrix} -\text{ant} \\ <-\text{cont}> \end{bmatrix} \rightarrow \begin{bmatrix} -\text{back} \\ +\text{cor} \\ +\text{strid} \\ <+\text{del.rel.}> \end{bmatrix} / __ \begin{bmatrix} -\text{cons} \\ -\text{back} \end{bmatrix}$$

The change [–back] links to the coronal marking convention, where the unmarked value is [+cor] in [–back,–ant] consonants (it is [–cor] in [+back,–ant] segments). Markedness rules are linked in sequence, so the immediate result of applying coronal markedness triggers a change in the value of del.rel. to plus (because of the changed value of coronal), and finally a change in stridency. To block this chain of secondary feature modifications and allow the output to be [k^{y} g^{y} x^{y}], the rule simply needs to explicitly specify that [coronal] is not changed:

$$(14)\quad [-\text{ant}] \rightarrow \begin{bmatrix} -\text{back} \\ -\text{cor} \end{bmatrix} / __ \begin{bmatrix} -\text{cons} \\ -\text{back} \end{bmatrix}$$

Because reassignment of the value of coronal is preempted with such a formulation, further changes to the segments do not arise. The added complexity of the latter rule predicts that [k^{y} g^{y} x^{y}] will be a less common form of velar palatalization. Stanley (1967: 404) similarly proposes that the output of any rule is subject to the segment structure rules of the language, so if a segment structure rule requires non-low back vowels to be round, then any rule inserting a non-low back vowel automatically undergoes the roundness redundancy rule.

Other limitations on rule operation were proposed, with researchers seeking a way to capture recurring and potentially universal generalizations while maintaining simple notation. An example of such a rule-external constraint is the Crossover Constraint (COC) (Howard 1972), which limits the interpretation of

variables in phonology.[12] Given the adoption of rule iteration, the star-parenthesis notation became superfluous, and was suspicious insofar as it was only used to express the notion "any number of possible rule foci." Elimination of the notation allowed a constraint on material appearing between the target (focus) and trigger (determinant) in a rule: "No segment may be matched with an element other than the focus or determinant of a rule if that segment meets the internal requirements of the focus of the rule."

The Crossover Constraint was seen as a constraint on string-to-rule matching, and not on possible rule statements. This allows a simple statement of the Menomini vowel raising rule with no mention of intervening features, which affects all long mid vowels and intentionally skips over all vowels, but extensionally does not skip long mid vowels:[13]

$$(15)\quad \begin{bmatrix} +\text{syl} \\ -\text{low} \\ +\text{long} \end{bmatrix} \rightarrow [+\text{high}] / __\ C_0\ (VC_0)_0\ C \begin{bmatrix} -\text{cons} \\ +\text{high} \end{bmatrix}$$

The effect "anything besides a long mid vowel" is determined by universal principle.

A related constraint is the Relevancy Condition (RC) (Jensen 1974):

> Only IRRELEVANT segments may intervene between focus and determinant in phonological rules. The class of segments defined by the features common to the input and determinant of a rule is the class of segments RELEVANT to that rule, provided at least one of the common features is a major class feature. If there is no common major class feature, then ALL segments are relevant.

This constraint operates in the context of a theory which (apparently) only had a generalized variable X and no infinite abbreviatory expressions. See Odden (1977, 1980), Jensen and Stong-Jensen (1979) for discussion.

Guerssel (1978) proposes the Adjacency-Identity Constraint (AIC):

> Given a string A_1A_2 where A_1=A_2, a rule alters the adjacency of A_1A_2 if and only if it alters the identity of A_1A_2.

The purpose of this constraint was to explain why certain rules did not affect geminate segments: for example, vowel epenthesis is blocked from splitting up geminate clusters.

Another constraint of the era, governing whether a rule could apply, was the Revised Alternation Condition (RAC) (Kiparsky 1973), a global constraint which states that "Non-automatic neutralization processes only apply to derived forms." The purpose of this constraint is to block application of rules such as assibilation in Finnish, which do not apply to lexical /ti/ sequences in [äiti] 'mother' but does apply to derived sequences, for example, [vesi] ← /vete/ 'water', [halusi] ← /halut+i/ 'wanted'.

The above constraints had "active" consequences, forcing a particular interpretation of the notation (interpretation of variables with RC, COC; causing non-specified changes via linking; blocking rule application with RAC and AIC), but some constraints simply state universal properties of rules. The Markedness Constraint of Houlihan and Iverson (1977: 61), conceptually related to SPE linking, requires that "Phonologically-conditioned neutralization rules convert relatively marked segments into relatively unmarked segments." Although their discussion does not provide explicit formulations of the rules under discussion, they do not suggest that this constraint results in any changes in how rules are stated or applied. Rather, this constraint expresses a well-formedness requirement on phonological grammars.[14]

Apart from such global constraints on rules which were held to be universal, holding of all languages and operations, there were also language-specific constraints applicable to single rules – that is, unformalizable conditions on rule application. One example is the blocking condition on Ojibwa T-palatalization (Kaye and Piggott 1973: 360, note). This rule is blocked when a sibilant follows, and Kaye and Piggott do not formalize the condition, stating "We are uncertain as to the formal status of this effect. It is our opinion that it does not form part of the T-Palatalization rule proper but rather is a condition ancillary to that rule." Kiparsky (1982b: 147) formulates English Trisyllabic Shortening as:

(16) $V \rightarrow [-long] \; / \; ___ \; C_0 \; V_i \; V_0 \; V_j$ where V_i is not metrically strong

Glover (1988: 225) formulates the epenthesis rule of Muscat Arabic as:

(17) $\varnothing \rightarrow V \quad / \; C_i \; __ \; C_j \;] \; (C_k V..)]_{Nominal}$
[+high]
Conditions: 1) Rule is optional when C_j is a fricative.
2) C_iC_j do not form a sonorant-obstruent sequence.
3) C_i is not identical to C_j.

Combined with the notion of phonotactic constraint, Newman (1968: 513) proposes the following schwa-deletion rule in Tera:

(18) ə → ∅ / ___ X (where X is not #)
Condition: Rule void where not permitted by phonotactic rules.

The relevant phonotactic rules (constraints) are that words are minimally CV and cannot end in a voiced obstruent or a cluster.[15]

Global rule conditions in general escaped formalization. For example, Kisseberth (1973) argues for a global condition on vowel shortening in Klamath which shortens long vowels either after $V{:}C_0$, or after two consonants when not followed by CV. This shortening only applies to long vowels deriving from vocalization of vowel+glide sequences, a condition which could not be formalized. Similarly, Miller (1975) posits a global condition on West Greenlandic assibilation, that it

changes /t/ to [s] after [i] before another vowel, but only when the preceding [i] is underlying /i/, not derived by epenthesis (epenthesis is shown to be ordered before assibilation) – this global condition is also not formalized. Thus despite best efforts, a number of factors determining rule applicability remained outside the scope of a fully formal theory of rules lacking recourse to plain-English restrictions.

The upshot of this section is that classical generative rule theory is characterized primarily by string-rewrite rules augmented by notational conventions referring to string classes, but there are also global limitations on the actions performed by rules – constraints. The constraint might trigger a repair (as in linking) or, more commonly, block rule application. The main characteristic of these constraints is that they are often held to be universal and global. As universals, the question of how these statements are formalized in a grammar need not arise, because the constraints are not part of a grammar. Such constraints were typically stated in prose (markedness and linking conventions were actually formalized). Most problematic were unformalizable language-specific constraints, undermining the concept "formal theory."

2.3 Evaluative Constraints

Evaluative constraints, as distinct from string-changing rules and ones guiding rule application, became particularly relevant in phonology via morpheme-structure conditions (Stanley 1967). The purpose of MSCs is to recognize redundancy in underlying forms: for example, English morpheme-initial nasals cannot be followed by a consonant. Previous work such as Halle (1959) would assume a zero specification for [consonantal] in the dictionary, and rules fill in a surface value.

Stanley showed how blank-specification undermined feature binarity and argued that phonology operates on fully-specified matrices, proposing that blank specifications be restricted to so-called "dictionary matrices." MS rules – rebranded as MS conditions (Stanley 1967: 424ff.) – are seen as statements of redundancy. Conditions either accept or reject matrices according to whether they satisfy or contradict the condition; conditions can be positive, negative, or "if-then." The latter type of constraint plays a significant rule in constraint-driven phonology, since it allows encoding cause and effect directly, for instance the if-then constraint "[–cons] ⊃ [+voice,+cont,–strid]" rules out vowels which are voiceless, stop, or strident, and fixes the locus of repair on the features *voice*, *cont*, *strid*, saying that vowels receive these three values.

Following research on derivational constraints pursued by Lakoff (1970, 1971), Kisseberth applies the notion of derivational constraint in his 1970 paper on phonological conspiracies. He argues that generalizations are missed in the standard account of Yawelmani:

> There are rather heavy constraints in Yawelmani phonetic representations on the clustering of consonants and of vowels. No vowel-vowel sequences are permitted. Words may neither end nor begin with consonant clusters. Nowhere in a word may more than two consonants occur in a sequence. (p. 294)

Specific rules of the language such as vowel epenthesis, syncope, consonant deletion, final Apocope, and a requirement on underlying representations regarding the presence of "protective vowels," appear to conspire to guarantee what we now understand to be the lack of branching onsets and rhymes. The formalism at the time provided no means of extracting this generalization and reducing it to notation. Note also that the concept "syllable" (restrictions on whose structure covers these generalizations quite simply) had no status in the formal theory at the time: see Goldsmith (this volume) for an overview of the syllable. Kisseberth argues that the evaluation metric *should* recognize the value of functionally related rules, even without the structural similarity required to bring them under the purview of abbreviatory notations. Fleshing out the formal details of the idea, especially how to express the notion "functional relatedness" is left for future research (p. 303). One part of Kisseberth's account posits derivational constraints against the sequences CCC, #CC and CC#. If a rule can only apply when "the output string would not be in violation of the derivational constraint," syncope can simply be stated as deleting a short vowel between two consonants, and the further fact that the consonants must be *single* consonants follows from the fact that if there were two consonants to the left or right, a forbidden triconsonantal sequence arises. Syncope can thus be simplified.

Shibatani (1973) argues for surface structure constraints, analogous to Stanley's MSCs. He argues that German must have a constraint requiring word-final obstruents to be unvoiced. This contrasts with the orthodox view that SSCs are redundant, because the facts which they cover are already explained by MSCs and the phonological rules of the language, wherefrom any SSCs can be deduced. Shibatani's argument emphasizes the fact of "independent psychological reality," the claim that SSCs are things that speakers "know" and therefore must be expressed as such in the grammar. Without a constraint to reflect the knowledge that *[bund] is not possible in German, speakers would have to not only look at existing words but also every imaginable word and apply the rules to these underlying forms to arrive at the conclusion that *[bund] would have no source in the language (could not come from /bundə/ or /bundö/): the procedure for evaluating "possible derivability" would be very complex.

Shibatani argues that rules are (often) redundant (p. 100) so the German devoicing rule [–sonorant] → [–voi] / __ ## is "identical" to his SPC2 which is stated as:

(19) IF: [–sonorant] ##
↓
THEN [–voiced]

Similar to SPE linking, Shibatani notes that SPCs encode repairs. Given the if-then format, he proposes "The convention further entails the imposition of the features given in the THEN-part of A/SPCs onto all the forms that meet the IF-part of the same constraints." The imposition of "then"-features localizes the constraint and the repair, so that both word-final regressive voicing agreement (20a) and progressive voicing agreement (20b) can be derived from SPCs.

(20) a. IF: [–sonorant] [–sonorant, αvoiced] ##
↓
THEN [αvoiced]

b. IF: [–sonorant, αvoiced] [–sonorant]##
↓
THEN [αvoiced]

This is a significant step towards elimination of rules, or at least demonstrating a significant level of interchangeability between rules and constraints.

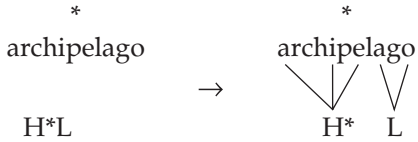
Sommerstein (1974) proposes a more formal implementation of Kisseberth's idea of conspiracy, where a rule can be marked to apply only if it improves harmony (in the sense of compliance with phonotactic statements) with respect to the motivating constraint, thus "A rule, or sub-case of a conspiracy, positively motivated by phonotactic constraint C does not apply unless its application will remove or alleviate a violation or violations of C" (p. 75). When connected to a constraint prohibiting final consonants, a language might delete final voiceless consonants and insert schwa after voiced consonants; under Sommerstein's proposal, these rules can be simplified by linking them to insertion and deletion rules, which can be stated as "delete voiceless" and "insert schwa," since phonotactically motivated rules only apply when violation of a constraint is alleviated. The choice of the word "alleviate" is also noteworthy: in his view, constraint satisfaction could be partial, presaging the OT view of relative harmony and gradient constraint violation.

2.4 Nonlinear Representations and Rules v. Constraints

The introduction of Autosegmental Phonology in Goldsmith (1976) gave substantial impetus to the expansion of constraints in phonology. Essential to Goldsmith's theory is the Well-formedness Condition, which states "All vowels are associated with at least one tone; all tones are associated with at least one vowel; Association lines do not cross." In line with the prevailing "trigger repairs when violations arise" viewpoint, Goldsmith states:

> Note that the Well-formedness Condition is in the indicative, not the imperative. A derivation containing a representation that violates the Well-formedness Condition is not thereby marked as ill-formed; rather, the condition is interpreted so as to change the representation minimally by addition or deletion of association lines so as to meet the Condition maximally. (p. 27)

The tune-to-text mapping in (21) occurs automatically, not by specific rule, to satisfy the WFC.

(21)


Similarly (p. 50), an argument for the autosegmental model "comes from the phenomenon of bidirectional spreading and, we would suggest, its *un*governed nature in these cases – that is, the spreading is not due to a specific phonological rule, but rather to the geometry of autosegmental representations, and its Well-formedness Condition. . . ."

The necessity of general conditions on structures, rather than explicit rules implementing an effect, is particularly compelling given the wide-spread phenomenon of tone preservation. As Goldsmith argues (p. 31), the alternative (proposed by Spa 1973) is a global rule that when a segment with H tone deletes, the H transfers to the nearest syllabic segment: but this constitutes a new formal object outside the purview of ordered rules. The significance of the WFCs is that they not only prevent certain relations, such as line-crossing but, like Shibatani's interpretation of SPCs, *demand* others, for instance that toneless vowels are prohibited – so if there is a toneless vowel, some tone must spread to that vowel. Goldsmith's account had a significantly lower dependence on explicit rule statements, and a higher use of representational possibilities interacting with general constraints.

The logic of autosegmental representation makes expansion of the role of constraints mandatory. In the theory of linear representations, especially with fully-specified underlying forms, it is easy to satisfy that aspect of the Naturalness Condition (Postal 1968: 61–62) which requires dictionary representations to map to some phonetic form without applying rules of a grammar – requiring only the application of universal conventions – because there was no such thing as a representation without an interpretation, all features being present in all segments.[16] Autosegmentalization meant that a representation might contain segments lacking a specification of voicing. For representations to be interpretable, and not simply due to the good graces of a particular rule but always interpretable in any language, universal conventions would be necessary to link up features or guarantee specifications when missing.

Research in the autosegmental paradigm was not univocal in seeking a shift in the direction of universal representational constraints. The version of autosegmental phonology proposed in Haraguchi (1975), also pursued by Clements and Ford (1979), Halle and Vergnaud (1982), and Pulleyblank (1986), depends more on language-specific rules to accomplish tone-to-TBU mapping. The first step in tonal mapping is an Initial Tone Association Rule; as characterized by Clements and Ford (1979: 181), "Initial tone association results from the application of rules which are language-specific, but drawn from a narrowly-defined set of rule schemata." The WFCs of Haraguchi and Clements and Ford are persistent and universal, and perform a 1-to-1 tone-to-vowel mapping. Other mappings are language-specific rules, so association of free tones to a vowel already bearing tone is mediated by a specific rule (Clements and Ford 1979: 191).

Halle and Vergnaud (1982) pursue an even more rule-driven account without a Well-formedness Condition, thus spread of linked tone (p. 73) is accomplished by a Mapping Rule applying only to free tones.[17] Because Halle and Vergnaud distinguish autosegmental versus "phonemic core" tone specification, where

autosegmental specification overrides core specification, autosegmentally toneless vowels are allowed on the surface without spreading being needed to fill in tone. Pulleyblank (1986) conjectures that the association conventions only apply at the beginning of a derivation, leaving derived 1-to-1 free tone/V configuration alone, to be repaired only by specific rule. Pulleyblank also relies on default specification as opposed to "phonemic core" specifications, but also dispenses with automatic spreading. Default specification has the flavor of both rules and constraints. Like markedness rules, default rules seem to be universal, but like rules, they "apply" at a particular point in the derivation – what that point is was a matter of discussion in underspecification theory; see especially Archangeli (1984).

The conflict between rule vs. convention-based grammar was also evident in syllable theory. The approach to syllabification in Kahn (1976) largely eschewed representational constraints. Syllabic analogs to Goldsmith's Well-formedness Conditions are proposed, so each [+syl] segment is associated with exactly one syllable, each [–syl] segment is associated with at least one syllable, and lines cannot cross. Kahn's tack, though, is to have explicit rules which achieve a well-formed state, hence he proposes Rule I, which states that [+syl] links to a syllable, and he does not appeal to action via universal convention. The language-specific rule-governed nature of Kahn's syllabification algorithm is especially made clear in his discussion of consonant clustering options in syllables:

> The system of rules assigning syllable structure to strings of segments, as envisioned here, does not refer back to some general set of constraints on possible word-initial and -final clusters which is pervasive throughout the phonology. It is rather in the syllable-structure assignment rules themselves that these constraints are found. Furthermore the constraints are not referred to by any other rules of the phonology. (p. 25)

The implicit assumption is that the syllabification rule would directly state that an onset could be [sp] but not [ksp]. Undermining this presumption is the fact that no rule was given to encode the restrictions. The reason why onset restrictions of English cannot reduce to well-formedness constraints in the sense employed in autosegmental tonology is that the required constraints are not universal across languages, or even within English – onset stop clusters which are not allowed via core syllabification (*[ptɛrədæktl̩]) arise in the output of later vowel deletion rules ([pətéɾo] → [ptéɾo]) "potato."[18]

Clements and Keyser (1983) pursue a more constraint-dependent approach to syllabification, exploiting positive conditions which license certain kinds of onset clusters (admitting [sp,st,sk] and [pl,pr,kr] etc. onsets), and negative conditions, which filter out a subset of positively licensed clusters (eliminating *[tl, pw] and various other more specific clusters). Rather than positing language-specific ordered rules to construct syllables, the Clements and Keyser approach posits general principles which are universal (with the parametric choices "delete syllable initial C" and "insert syllable final C" as well as allowing sequences of vowels and consonants, pp. 28–30), persistent, and which interact with language-specific

admissibility conditions. A general "grouping" process is constrained (p. 37) by the Onset First Principle:

(a) syllable-initial consonants are maximized to the extent possible with the syllable-structure conditions of the language in question,
(b) Subsequently, syllable-final consonants are maximized to the extent consistent with the syllable-structure conditions of the language in question.

Onset First has a mixed status qua principle/constraint versus rule. It is termed a principle, and it is not subject to the standard linear ordering requirements of rules – but, Clements and Keyser also refer to these principles as rules. For example, (p. 54) the Resyllabification Convention states "The output of every rule is resyllabified according to the syllable-structure rules examined up to that point in the derivation," which asserts that these are rules, but attributes to them a property of constraints, namely everywhere-applicability. One rule-like property of syllabification is that it can have a derivational "endpoint," that is, at a certain step in derivations, it ceases to function (p. 55): "We propose, then, that individual grammars may specify a point in the set of ordered rules at which the Resyllabification Convention becomes inoperative . . ."

In addition to the aforementioned principles governing basic autosegmental associations, the repertoire of constraints includes the Twin Sister convention prohibiting adjacent identical feature values on a single feature-bearing unit (Clements and Keyser 1983), the Linking constraint (Hayes 1986: 331) which states "Association lines in structural descriptions are interpreted as exhaustive"; the Shared Features Convention (Steriade 1982) which forces merger of identical feature values under certain conditions, in response to the application of a rule. Other constraints on phonology were widely employed in this era, such as the Strict Cycle Condition, the notion of structure preservation, and the ideas of structure-building versus structure-changing rules, especially the related notions feature-changing vs. filling, which allowed rule theory to avoid explicit reference to zero.

While the role of independent constraints and interpretive conventions expanded considerably in the autosegmental era, attention was also paid to the theory of rule formulation. Pulleyblank (1983: 55–56) advances a standard symbolic notation for expressing rules operations where in addition to the notational standards introduced in Goldsmith (1976b), a line to Y means "is linked,"[19] a circle around Y means "is not linked" and a line to Y followed by a line from Y to "x" means "rightmost," that is, "a link not followed by another link." The notion of features being organized into constituents (Clements 1985; Sagey 1986) made possible the single-node characterization of rules (Clements 1985: 244), "assimilation processes only involve single nodes in tree structure," or more generally, rules operate on only one object. McCarthy and Prince (1981: 1) claim "a rule may fix on one specified element and examine a structurally adjacent element and no other," limiting the class of well-formed rules significantly (a proposal in part made plausible by expanding the class of "elements").

Discussion of constraints and rules in the context of autosegmental phonology would be incomplete without mention of the Obligatory Contour Principle: see Leben (1973), Goldsmith (1976), Singler (1980), McCarthy (1986), Odden (1986, 1988), and Yip (1988) inter alia. The basic statement of the OCP is "Adjacent identical elements are prohibited," thus two adjacent H tones would be prohibited, two adjacent identical specifications for voicing would be prohibited; generalizing the original version of the OCP somewhat, the principle could also rule out adjacent identical segments (identity with respect to the whole set of features) or adjacent homorganic segments (identity with respect to a subset of features).

One view of the OCP, advanced in McCarthy (1986), is that it is an absolute representational universal; a competing view set forth in Odden (1986, 1988) is that it is not directly part of linguistic theory but is a formal accident resulting from an interaction between language learning and representational theory, and is only formally instantiated as a language-specific rule. The problem for the representational universal view is that the highly variable naure of the OCP – its effect, viz. limiting underlying contrasts, triggering a process, blocking a process; what unit it applies to (tones, place of articulation, major articulators only, laryngeal features, whole sets of features); the degree of adjacency (strictly adjacent segments, in adjacent syllables, within the same word); even whether it is obeyed or simply ignored. All of these considerations point away from the idea of a hard universal, analogous to the No-Crossing constraint. Such parochiality was typically seen as evidence for rule status whereas the concept "constraint" was traditionally reserved for hard universals; nevertheless, within Optimality Theory, the OCP, once joined with "constraint family" and "violable constraint," remains a universal constraint. Reiss (2003) on the other hand draws a different formal conclusion, that rule theory requires variables, quantifiers, and equality computations.[20] As Reiss points out, homorganic syncope ("anti-antigemination") in Yapese and Koya, where C_iVC_i and only C_iVC_i syncopates to C_iC_i, cannot be explained by appeal to sharing of a place node, so some reference to identity of values is needed, and OCP effects are easily subsumed under a general theory that includes feature identity.[21] See also Baković (2005) for related quantificational analysis within OT.

To summarize the course of rule theory, there has been a steady progression of ideas, from minimal reliance on the guiding hand of UG and more emphasis on explicit statement of directly interpreted operations, to a greater reliance on conditions, some language specific but in the mainstream view universal, which are stated independent of the rules that control derivations. The main difficulty facing the theory is the assumption that rules are language-specific whereas constraints (conditions, principles) are true of rules in general, and yet hard and fast constraints turned out to be difficult to come by. Many putative conditions required specific assumptions about representations which were highly controversial. The OCP debate highlights both of these problems, in that manifestations of the OCP are sufficiently common across languages that it cannot be dismissed as a coincidence; and yet it is not an absolute representational universal.

3 Parametric Rules

If rules are non-recurring parochial statements of the mapping from input to output, and grammars contain just rules and representations which rules act on, then generalizations which recur within a language or across languages are expected only to the extent that they might arise more than once by random combinations of symbols into rules, according to a theory of rule formalism. Thus we expect rules of regressive nasal assimilation or voicing assimilation because such rules are possible in rule theory, and we do not expect the mapping /p,l,i/ → [r,o,t] / __[s,p,e] which is not a rule in phonetic feature-based rule theory. But within possible rules, there is a significant disparity between observed frequency of rules and their combinatoric probability, given the free combination of elements according to a syntax of rule formulation. Nasal place assimilation and post-nasal voicing are common, but post-nasal devoicing and "continuancy assimilation" (e.g. /xt/ → [kt], /ks/ → [xs]) are extremely rare and possibly non-existent. Formally speaking, there is no basis for this, since continuancy assimilation is the same operation as place assimilation, simply applied to a different node in the representation, and post-nasal devoicing is expressible as dissimilatory delinking, a known process in language. The frequency of consonantal homorganicity conditions on rules and the rarity of analogous laryngeal identity conditions cannot be formally explained just on the basis of formal properties of rules. From OCP investigations, we know that these tendencies cannot be hard universals – there is no absolute requirement that nasals *must always* agree in place with the following consonant.

The idea of a "parameter" is well suited to resolve recurrency with violability. A parameter is a fixed choice given by UG, which narrows the degree of freedom to less than that given by free combination of symbols, but still provides a degree of freedom greater than zero. The notion of "parameter" is introduced in Chomsky (1964: 315), who states:

> Even if conditions are language- or rule-particular, there are limits to the possible diversity of grammar. Thus, such conditions can be regarded as parameters that have to be fixed (for the language, or for particular rules, in the worst case), in language learning . . . It has often been supposed that conditions on application of rules must be quite general, even universal, to be significant, but that need not be the case if establishing a "parametric" condition permits us to substantially reduce the class of possible rules.

This approach particularly flourished in syntax in work emanating from Chomsky (1981).

Parameter-like theories of phonology were pursued in Natural Phonology (Stampe 1972), which posits that humans are endowed with a list of innate, substantive phonological processes, some of which must be suppressed in the course of language acquisition. The theory of Atomic Phonology posits that the core of phonological systems is a collection of given basic processes (such as palatalization,

vowel nasalization, etc.), termed "atomic rules," and "complements"[22] of the atomic rules present in particular grammars. Dinnsen (1979: 31) thus posits that "The theory of atomic phonology maintains that all linguistic variation requiring distinctly varied formulation of phonological rules is predictable from a set of atomic rules and universal principles of grammar." The atomic rule of final devoicing is:

(22) $\begin{bmatrix} -\text{son} \\ -\text{cont} \end{bmatrix} \rightarrow [-\text{voice}] \; / \; __ \; \#$

Any language with final devoicing must have *at least* this form of the rule – a language with just "velar stop devoicing" or "fricative devoicing" would be impossible since those rules are (by hypothesis) not atomic rules. Restricted sets of options are made universally available, but they may be overridden – the process may be suppressed, the atomic rule may not be selected, or a complement rule may be selected, as long as the atomic rule is.

Non-linear phonology in the 1980s also saw an increased reliance on rule construction with formal and substantive parameters. As discussed above, core syllabification in Clements and Keyser (1983) invoked consonant insertion and deletion parameters. Hayes (1980) proposes "that the characteristic stress rules which occur in language after language are all derivable using a fairly simple rule schema, in which a number of parameters may be set independently of one another." While arguing for an absolute, inviolable universal interpretation of the OCP, McCarthy (1986: 256) also allows that "The alternative and, I think, the best way to account for any nonuniversality in the OCP, if clear violations arise that are not susceptible to reanalysis, is to consider the OCP a parameter of Universal Grammar whose unmarked value is 'on'." Substantial use of OCP parameters in rules is found in Yip (1988); parameters play a major role in certain typological studies, such as Cho (1990) for consonant assimilation and Hayes (1995a) for stress. Universally fixed choices for adjacency conditions are discussed in Odden (1996), and Calabrese (1988) proposes a rule-based theory augmented with parametrically-selected negative filters on segments and universal clean-up rules.

The fundamental work in parametric rule theory is Archangeli and Pulleyblank (1994), which articulates a general parametric theory of rules, Grounded Phonology, combining absolute conditions and universal choices for rule formulation. See also Davis (1995) for an application of the theory to rule statement in Palestinian Arabic. The concept of a constraint or condition is strong, according to Archangeli and Pulleyblank (1994: 14): "Wellformedness encodes the requirement that no representation may be allowed, even temporarily, to violate conditions," thus (p. 14) "Given an input representation of a particular type, a *convention* predicts a single related output representation." The theory defines rules in terms of conditions in fixed boxes including *function, type, direction, iteration*.

Rules specify an argument (the focus in traditional terminology), so a rule spreading [–ATR] has the argument [–ATR], and can have structure requirements on argument and target (whether they must be unassociated or not). Finally, rules have an "other requirements" box for context conditions, such as whether

a rule applies only to certain morphemes or string properties like "word finally if preceded by L," or the Menomini target condition μμ (p. 379) which states that only long vowels can undergo ATR harmony. This much of the theory essentially re-states aspects of standard non-linear rule theory.

The most significant difference from standard rule theory resides in the substantive "grounding conditions" on argument and target, which specify phonetically motivated if-then relations between features in a path. An example is ATR/LO, which states that if a vowel is [+ATR], it should be [–low]. Imposed as a target condition, only non-low vowels could undergo an ATR spreading rule (a common restriction on ATR spread), and as an argument condition it states that [ATR] spreads only from a non-low vowel. The theory presumes specific sets of feature relations, thus six $\alpha F_i \supset \beta F_j$ relationships between ±ATR, ±hi and ±low are postulated as exhausting the range of phonetically grounded conditions. This listing precludes combinatorially possible relations such as "if [+hi] then [–ATR]"; evidence for such a condition would be a case where [–ATR] spreads only to a high vowel or from a high vowel (see Poliquin 2006 for an example from Canadian French).

The question will naturally arise whether there is a substantial difference between the parametric rule approach of Grounded Phonology and similar works, and the Principles and Parameters account of Halle and Vergnaud (1987), or Paradis discussed in the next section. Hayes (1995: 55) aptly characterizes the matter as follows:

> An interesting problem within parametric metrical theory is to what extent the parameters characterize rules versus grammars. Here, we will conservatively assume that parameters characterize rules. However, the possibility that they have more general scope, as suggested by HV [Halle and Vergnaud 1987: DO], is an appealing one: for example, it predicts that when more than one rule creates feet, the feet created should be the same.

A pure-parameter approach would say that the scope of the parameter is the particular representational object, and the prediction is that a language should not have multiple rules spreading or deleting a given feature, except if a parameter holds only of one lexical level. In a parametric rule approach, the scope of a parameter is the given rule, which allows more than one rule focusing on a particular feature such as nasal or H tone. Odden (1981) argues that Karanga has over a half-dozen each of partially similar rules raising L after H and lowering H after H, differentiated by subtle contextual properties, but perhaps with a highly articulated theory of level, these rules could be reduced to single parameter settings.

4 Constraints-Only

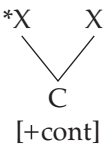
Given the expanding role of constraints in phonology, it would seem a hindsight-obvious simplifying move to attempt a theory without rules. As discussed in

Section 5, it is unclear which theories are "constraint-only" since there is no clear characterization of "constraint" as distinct from "rule," and naming conventions are variable, for example Karttunen's (1993) paper is about "Finite State Constraints," but also talks about these constraints as "rules." A characteristic of the rule-based approaches of the preceding section is positing that the engine underlying phonology is a set of string-changing rules. Constraint-based theories deny this, and may deny that there is any string changing at all (Declarative Phonology) or view string changes as automatic responses to representational requirements (TCRS and OT). This section considers four approaches that can reasonably be considered constraint-only: TCRS and similar Principles and Parameters (P&P) theories; the fixed-level approaches of Goldsmith, Lakoff and Karttunen; Declarative Phonology; and OT.

4.1 Principles and Parameters

The main representative of P&P phonology is the Theory of Constraints and Repair Strategies (TCRS), articulated in Paradis (1987, 1988), building on work by Singh (1987) and Piggott and Singh (1985). The essential difference between strict P&P phonology and parametric rules is that the latter theory has a linearly orderable grammatical object, but the P&P approach only states conditions on representations, and derivational steps are given automatically by the theory.

Paradis (1987, 1988) argues for a repair driven model, based on the contention that phonological rules are "contextual and arbitrary" but repair strategies are context-free and "motivated," the context and motivation of the repair being found in the constraints. In TCRS, constraints can be universal ("principles") or language-specific. Examples of presumed universals are the OCP and Prosodic Licensing (all units must belong to higher prosodic structure). Constraints have either a blocking effect or, if blocking is impossible, they trigger a repair – "insert," "delete," or "change." The theory has a number of particular parameters, such as "Spread Nasal," so if "Spread Nasal" is set "on" for a language, then nasal must spread. Other parameters accept/reject particular sequences, for example sequences of non-high vowels may be accepted in some languages, but are generally rejected due to a parameter setting. Parameters may also be set according to lexical phonology domain (similar to Goldsmith's Harmonic approach). An important feature of Paradis' constraints (found also in Shibatani's account) is that constraints have a focus. The Fula constraint:

(23) 

has a segmental focus on the feature [+cont]. The locus of a repair would be that feature, and in Fula, would-be geminate continuants change to stops (rather than degeminating).

Paradis also points to the "many effects from one constraint" argument made by conspiracy theorists, an example in Guere being the constraint against non-high vowel sequences, which not only limits underlying forms (there are no non-high vowel sequences within morphemes – an MSC), but it also triggers vowel raising and vowel deletion. Similarly, OCP-labial causes both hardening (*wʊ* → *gʊ*) and deletion (*kwu* → *ku*).

TCRS separates morphophonology and automatic phonology. Paradis (1988: 5) notes "I do not claim, then, that there are no rules but rather that these are morphologically conditioned processes." It is not clear whether such processes would be in a separate grammatical module, though the disposition of morphophonology in Singh's theory (1987: 282) is clearer: it "cannot work without giving up what has seemed to be the non-negotiable heart of generative phonology: the assumption that even non-automatic morphophonology is a part of phonology." An example of a phonological rule consigned to morphology mentioned by Singh is English Trisyllabic Laxing (accounting for the alternation *serene ~ serenity*). This would be an example of how precluding classes of phenomena may allow a formally more constrained theory, while making comparison of theories (standard rule theory versus P&P phonology) meaningless because they are theories about different things.

4.2 *Fixed-level Accounts*

Another approach to stating phonological regularities, arising from work in computational linguistics, especially Koskenniemi (1983), relies on directly stating relationships between input and output (or some similar fixed set of levels). One of the main concerns of fixed-level approaches is elimination of extrinsic rule ordering, also a goal of the Unordered Rule Hypothesis (URH: Koutsoudas, Sanders, and Noll 1974). Early implementations of unordered rules failed because the claim of persistent reapplication of rules was falsified by counter-feeding relationships, and direct mapping theories (Kenstowicz and Kisseberth 1977 291 ff.) were falsified[23] by feeding relationships. The problem for direct mapping is that if rules can only refer to what is present in underlying forms, cases such as Lardil are impossible to express without redundant recapitulation of the conditions in Apocope. In the Lardil derivation /tjumputjumpu/ → *tjumputjump* → *tjumputju*, where Apocope feeds non-apical deletion, the conditions for consonant deletion are not present in the underlying form and, in light of the fact that /kuŋka/ → [kuŋka] without Apocope (because of a word-minimality restriction), an elaboration of non-apical deletion which allowed deletion of intervocalic consonants must repeat the conditions for Apocope.

In the version of Karttunen (1993) (see also Karttunen, Koskenniemi, and Kaplan 1987), a phonology is modeled as correspondences between input and output. Rather than producing the required output from a set of rules which modify an input, the constraints accept (or reject) pre-existing pairings of input and output, based on the properties of the input and output – which means that the constraint has simultaneous access to the input and the output (not possible under the

Markovian conception of rule). In Karttunen's notation, "u:" means "lexical u," ":u" means "surface u," and "⇔" expresses the input/output relation "is realized as . . . in the context . . . and nowhere else." Examples from Finnish which are relevant in accounting for the mapping in Finnish *kaNpan:kamman* are as follows.[24]

(24) N:m ⇔ ___p:
p:m ⇔ :m __

This means "input *N* is realized as *m* only before underlying /p/" and "input *p* is realized as *m* only after surface *m*." For the problem of Lardil Apocope and non-apical deletion, non-apical deletion could be stated as:

(25) [non-apical]:∅ ⇔ __:∅* :#

that is, an input non-apical must map to an output null before an output word boundary, disregarding deleted segments.

An alternative graphic representation of these relations is adopted by Lakoff (1993) and Goldsmith (1993a), who recognize three levels of structure,[25] the M(orphophonemic), W(ord) and P(honetic) levels. These levels describe respectively a description of phonological properties of the morpheme, the word (with minimal redundant information), and the phonetic output. Goldsmith's Harmonic Phonology constraints (Goldsmith 1993a) for vowel lowering and Apocope in Lardil are as follows:

(26)
M	[V]]$_{word}$		M	VCVCV]]$_{word}$
↕	|		↕	|
W	[-hi]		W	∅

Goldsmith's Harmonic Phonology addresses the well-known rule ordering relationships, by distinguishing intralevel and cross-level rules. Intralevel rules are held to be "harmonic," that is, the string is modified to the point that no further increase in harmony (satisfaction of target condition) results, thus *tjumputjump* loses final consonants until the perfectly harmonic string [tjumputju] results. This is analogous to the repeated application of rules in the URH which allowed feeding relations (but was falsified because of the existence of counter-feeding).[26] Cross-level rules, on the other hand, can be harmonic or non-harmonic, the latter meaning that there is a single evaluation of the relationship between levels. The relationship illustrated by the mapping /tjumputjumpu/ → [tjumputju] with respect to vowel lowering and Apocope (*[tjumputja], *[tjumpu]) exemplifies how cross-level rules can accommodate counter-feeding as a function of the rule itself: the rule only demands that a vowel which is word-final at the M-level correspond to zero (or a non-high vowel) at the W-level, and the last vowel of [tjumputju] is not word-final at the M-level. With three levels, five classes of rules are defined (three which describe properties of representations at the level and two which describe the relationship between adjacent levels), and the empirical claim is that this suffices to handling all rule ordering effects.

4.3 Declarative Phonology

Declarative phonology explicitly shares theoretical assumptions with the declarative syntactic theories HPSG (Pollard and Sag 1994) and LFG (Bresnan 1982), and like HPSG is, according to Bird, Coleman, Pierrehumbert and Scobbie (1992: 1), "an attempt to do away with the ordered derivations and the concomitant feature-changing rules of traditional generative phonology." The declarative paradigm is committed to non-algorithmically describing static properties of linguistic strings by continuous elaboration of a description where, as characterized by Levine and Meurers (2006: 377), "all representations which play a role in licensing a particular string are simultaneously and completely part of the model of the linguistic object being licensed." All statements in a declarative account must be true, that is, there can be no exceptions from any source to rules, and constraints cannot be violated.

The DP view is that a phonological representation is a "description of a class of utterances" (Bird and Klein 1994: 456), which refers to a narrower class of utterances when the description is more fully articulated, or a broader class of utterances when it is less fully articulated. For instance, English *p* can be described without mentioning aspiration or glottalization, in which case all *p*s would be subsumed under that description, or it could be described as "aspirated," in which case only the syllable-initial ones are being described. Descriptions of linguistic objects are said to be partial in that they do not specify every detail of an utterance – they are descriptions of *classes* of utterances, so the details distinguishing one utterance from another within the class will not be part of the class description. Questions of formal representation become paramount in a declarative phonology, and representations can be rather complex. Other examples of DP research are Bird (1995), Scobbie (1997), Coleman (1998, 2006), and Hoehle (1999). A very similar partial-description approach to phonology, relying on the notion of property-percolation and eliminating all feature-changing in favor of lexical allomorph selection is proposed in what appears to be the first generative constraint-only theory, Guerssel (1979), though DP does not appear to a have been influenced by that work.

Allophony is straightforward for DP, which like American Structuralism sees the phoneme as a descriptive device for subsuming a class of phonetic realizations. Little information is available on how DP treats neutralizing processes, which pose a problem for the non-destructiveness requirement of the theory. The German root-object meaning "federal" manifested in attributive *bund-ə* must be distinct from the root-object meaning "colorful" manifested in attributive *bunt-ə*, but the two root-objects are pronounced the same in uninflected [bunt] "federal; colorful." The standard feature-changing account is impossible since underlying information would not be present in all instances of the object being modeled (the root "federal"). Based on analyses by Bird in Bird (1995) and especially Bird, Coleman, Pierrehumbert and Scobbie (1992), it seems that this problem could be reduced to disjunctive allomorph selection as practiced by Trubetzkoy, Item-and-Arrangement morphologists (Hockett 1954) and Natural Generative Phonology (Hudson 1975; Hooper 1976). The representation would be enriched so that [bunt] "federal" could be /bunT{+voice,∅}/ and [bunt] "colorful" could

be /bunT{–voice,∅}/. Another approach to the problem is to deny the existence of neutralizations.[27]

4.4 Optimality Theory

While Optimality Theory relies on constraints, as McCarthy (2002) points out, it is a theory of constraint *interaction*, not constraint substance. OT has no necessary position on whether there is a constraint Onset requiring syllables to have onsets, and no necessary position on banning a constraint Coda obligating syllables to have a coda, or *Onset prohibiting onsets (p. 46, Note 13).

An OT constraint is a requirement which should be true for forms, but unlike most constraint-based theories, violations of constraints are possible, indeed unavoidable. For any violation a mark is assigned, and the output of the system is determined based on a computation over the set of marks. It is hard to say what syntactic form constraints have, since the theory does not hold to any particular idea of the syntax of constraints. Rules were held to be constructed on the basis of experience, using universally-defined primitives combined according to a particular syntax of rules. OT constraints are claimed to be entirely universal, so it would make little sense to talk of "constructing" constraints according to systematic principles. The fact that there is a constraint *NC̥ does not imply that there is also a constraint *NZ (although there probably has to be such a constraint in OT), *SC̥, *N[sonorant], or any other pairing of two consonants. OT constraints being universal, it would be difficult to pin down matters of actual form, since they are invariant across languages. In plain English, the Onset constraint can be stated positively as "A syllable begins with an Onset," or negatively as "no syllable may begin with a vowel." Constraints can also be stated symbolically, in which case they are usually stated negatively – *NC̥ or *[$_{\sigma}$V. Often constraints are simply named, for example, *Complex, when the function of the constraint is presumably obvious. As in Atomic Phonology, the set of constraints needs to be discovered.

Nevertheless, systematic aspects of constraints have been proposed, in the form of constraint schemata. For instance, OCP seems to represent a family of constraints, probably applicable to any feature or node; there seems to be a class of related constraints on identity with different adjacency requirements (see, for example, Bickmore 2000 for distinct rankings of adjacent vs. general Uniformity violations), or relativization to different morphosyntactic levels such as "stem" or "word"; constraints can subdivide into various positional versions. Another systematic form of constraint is the class of alignment constraints, which follow a general formally defined template Align (Edge, Category, Edge, Category).

In OT, constraints are ordered (ranked) and violable – ranking is relevant only to regulate conflicts arising from the impossibility of satisfying every constraint, and ranking makes violability possible. Unlike TCRS or DP, constraints can be violated. OT constraints somewhat resemble parameters in TCRS, which are not enforced in all languages but are potentially available in all languages, thus

allowing unenforced universals. However, in TCRS, if a parameter is set "on," then it is on and enforced throughout the language (or, the lexical level where it is on). The entire content of a grammar is the ranking of these constraints.

The original version of OT, without faithfulness conditions and obeying Containment – "the input is literally contained in the output, with no losses" (Prince and Smolensky 1993: 111) – was effectively monostratal, that is, constraint violation was determined only by inspecting the properties of the output. The advent of Correspondence theory moved OT into the realm of being at least a two-level theory like Kimmo morphology, since constraint satisfaction required inspection of both the input and the candidate itself. The advent of inter-candidate correspondence (Output-Output constraints and Sympathy constraints, inter alia) continues the representational enrichment of OT. In a recent development in OT, the theory of candidate chains (McCarthy 2007), the mapping in Yawelmani from /c'u:m-hin/ to [c'omhun] involves selecting the winning candidate which is a chain of forms <c'u:mhin, c'u:mhun, c'o:mhun, c'omhun>, having as many virtual levels as steps in a rule-based derivation.

5 Interchangeability of Constraints and Rules

The main difficulty in deciding between rules and constraints as the best model of language is the varying metaphysical implications (but not necessarily entailments) of these concepts. Does the concept "rule" entail a real physical operation in time; is a constraint a Platonic requirement that is instantaneously "somehow true?" Is it a disadvantage for a theory to have "productions?" Is it meaningful for a theory to talk as though forms already exist, waiting to be evaluated?

Linguists have, to a considerable extent, been willing to set aside strong commitment to particular metaphysical interpretations of theories, and disputes tend to center on the weak generative capacity of theories. The idea of interchangeability of methods has a venerable tradition in generative grammar, owing in no small part to the results of Chomsky and Miller (1958). The Chomsky hierarchy of production rules in formal language theory has a mathematical equivalence to automata said to "accept" certain languages, whereby Turing machines accept Type 0 languages (the languages produced by unrestricted rewrite systems), linear bounded automata accept context-sensitive languages, pushdown automata accept context-free languages, and finite state automata accept regular languages. A grammar producing a given class of strings is weakly equivalent to some machine that only accepts that class of string and rejects all others. This fact gives rise to the appearance of interchangeability of rules and constraints.

McCawley (1968) advocates a non-production oriented interpretation of base rules, which are understood to state admissible mother-daughter node relations where NP may dominate *det* and *N*, rather than stating how the object NP is converted into a sequence of objects, *det* and *N*. Lakoff (1970: 627–628) provides an insight into the notions "constraint" and "rule," suggesting that they are not very different objects.

> phrase-structure and transformational rules . . . are local; they define well-formedness conditions on individual phrase-markers and on pairs of successive phrase-markers in a derivation . . . *Transformations are essentially local derivational constraints, in that they filter out those pairs of successive trees which are transformationally related from those which are not.* (emphasis added)

Dinnsen (1972: 2) similarly notes that "phonological rules thus establish grammatical relationships between adjacent lines in a derivation."

Stanley (1967: 393) states that ". . . a morpheme structure rule can be interpreted both as a statement of a constraint on phoneme sequences and as an algorithm for predicting redundant feature values in phoneme sequences. The morpheme-structure rule itself is neutral as regards its interpretation." This again points to the recurring observation that string-rewrite rules and string-evaluation statements may be notational variants, from the perspective of the classes of strings that they describe. Finally, in comparing parallel and sequential descriptions, Karttunen (1993: 174) says "One important lesson that has been learned about the two styles of description is that in phonology they are formally equivalent." Thus the Finnish constraint regulating the *p* ~ *m* alternation stated in Two-Level Morphology as "p:m ⇔ :m __" means "accept an input *p* matched with output *m* just in case output *m* precedes," the rule "p→m/ m__" means "change *p* into *m* when *m* precedes," and the input-output relationships are the same whether you interpret the generalization as a well-formedness constraint or a production rule.

Translation between OT constraints and production rules may be straightforward,[28] since the proposition asserted by a constraint has an analog to some aspect of a production rule. The rule [+syl] → [+hi]/ __[+nasal] $_{\sigma}$] can be re-expressed as well-formedness constraints addressing the structural description, such as *[+syl,–hi] [+nasal] $_{\sigma}$], with limitations on repair strategies via faithfulness and markedness referring to the complement of the changing features (e.g. Faith(+syl), Faith(nas), which prevent denasalization or resyllabification as repairs) – in general, keep all things the same, except that which changes. It is very likely that such a translation could be automated, though hand-coding the markedness and faithfulness relations could lead to a more streamlined characterization of the constraints relevant to a process, just as hand-coding the composition of rules in two-level phonology can lead to simpler sets of regular expressions.

In short, if we are committed to neutrality as to metaphysical interpretation (and we are *not* all committed to such neutrality), a theory describing language as a system of operations replacing objects with other objects in real time is extensionally indistinguishable from a theory describing language in terms of separating wheat from chaff, in a pre-existing set of language objects. It is more productive to focus on properties of specific rule versus constraint theories and ask, which properties do phonological systems have? Some property can always be identified as defining the "line in the sand," sacrificing other considerations in its defense. As is well known, there is a trade-off between statement-simplicity and ordering – simplicity can be purchased at the cost of imposing order on processes (either derivational rule ordering or constraint ranking). Is simplicity and generality of statement so important that ordering is tolerable? Put the other

way, is ordering so repugnant that massive rule complication is actually preferable? Why is no-ordering intrinsically superior to simplicity (or the converse)? How much complication is needed to avoid derivational steps and ordering? Simplicity of the metatheory itself is also a consideration in theory selection. Is there validity in the Occam's Razor argument that constraints-only is conceptually simpler than rules plus constraints? Is that simplicity negated by the fact that OT also requires adding Gen and Eval algorithms to grammar? Because such philosophical questions are hard to answer decisively, it may be more fruitful to look for empirical answers, but even then, compelling and unequivocal evidence is hard to find.

5.1 Globality

A supposed difference between rules and constraints is globality and the conspiratorial nature of constraints. Rules are linearly ordered and constraints are classically unordered requirements on representations having numerous sources, so finding conspiracies would seem to support constraints over rules. However, it has long been known that rule ordering is not strictly linear because of the cycle; furthermore, it has been proposed in the rule-based context (Chafe 1968; Halle and Vergnaud 1987) that there are unordered "persistent rules." Some constraint theories have ordered domains (Paradis 1988; Kiparsky 2008a), and constraints in Harmonic Phonology fall into three ordered levels with two rule-governed transitions. P&P phonology allows constraints to have a "cutoff" within the derivation, and insofar as parametric rule theories such as Grounded Phonology are at heart a fusion of the notions "independent constraint" and "particular rules," constraints can be quite localized. Thus conspiracies do not automatically argue for constraints over rules, any more than "opacity" automatically argues for rules over constraints.

Since constraint and rule theories have resources for expressing globality, the productive question to raise is, what kinds of properties seem to be global? Rules and constraints alike operate on possible representations and either modify them or say something about them (whether they are "allowed"). Language is a system where symbols can be defined in terms of other things (symbols or perhaps physical properties). A universal syntax of representations for defining these symbols therefore establishes a baseline of globality: if an imaginable combination of primitives is not within the scope of linguistic representation – the would-be symbol is undefined – then rules will obviously show the effect of that fact. It is a fairly well supported hypothesis, at least in generative theories, that languages include (27a) but not (27b).

(27) a.

b.

The need for representational constraints on defined linguistic objects can hardly be questioned – there are feet, and they have a specifically defined nature which then limits (constrains) what a foot can be: but it is debatable what those representational objects are, and whether definitions are universal or can also be language-specific. The vowel objects [æ] and [ɑ] are both defined (generated) and used in English and Finnish, but the object [æ] is not defined in Italian. Global structure-preserving effects which limit the operation of rules to only deal with defined symbols would be one source of global effects in rule theory. Another kind of symbolic object that requires definition would be a prosodic constituent – onset, coda, foot. If "onset" is defined in some language as "sonority-decreasing sequence of consonants," then a sequence *rp* will not be an onset, and as long as phonological rules are restricted to producing defined outputs (and assuming that prosodic licensing is a requirement of representations), conspiracies regarding constituent makeup are expected.

A requirement for creating only defined objects is expected to result in conspiracies – segmental and prosodic structure-preservation effects – but simple concatenation of objects would not, by the same logic, lead to conspiracies. A linear segment sequence coronal+labial might be precluded by a constraint, and such a limit would not be due to the definition of a specific segmental or prosodic object. Multiple references to the exclusion of such a sequence in a rule system could then be an argument for the notion of "conspiracy" extrinsic to the system of rules, and thus an argument for the autonomous constraint. A case for including constraints in the theory of grammar would come from a language with multiple rules defined on concatenations of segments, having an evidently unified teleology but a disunified collection of structural changes, at least presuming that the similarity must be captured in the grammar. One example could be the fact that in Karanga, over a half-dozen rules lowering H tone to L after H are motivated, and given the various subtle differences in morphosyntactic requirements and other phonological conditions on the rules, it cannot be maintained that there is just one rule H→L/H__.

A second type of conspiracy, one that is fairly widely attested, is the grammar with multiple rules eliminating vowel sequences – glide formation, vowel deletion, and vowel fusions. Odden (1996) analyzes Kimatuumbi's vowel hiatus-resolution processes in terms of six specific rules, without explicit encoding of a common teleology behind these rules in the grammar. The constraint-based criticism might be that this leaves uncaptured a unifying generalization expressed through a motivating constraint against V-V sequences. The rule-based response would be that this is not a generalization needing to be captured in the grammar, and that expressing the generalization via a handful of separate rules is appropriate, since V-V hiatus resolution is fairly idiosyncratic and does not reflect a general fact of Kimatuumbi which, unlike Luganda, is rather tolerant of vowel hiatus, which is resolved in only around half of the contexts where it arises. The methodological question underlying the conspiracy argument is whether a grammar should directly encode all imaginable descriptive generalizations about the language. Just as it would be invalid to argue against a constraint-based account of phonological

processes on the grounds that it requires multiple constraints to fully express them, it is also invalid to argue against rule-based phonology by *presuming* that a grammar must contain single constraints that directly state teleological goals and then criticizing rule-based grammar for not having constraints.

5.2 Derivation-like Properties

Two main characteristics give a "derivational" character to theories, namely time invariance and multiple representations. Time invariance is a concept from signal processing, where a system is time-invariant in case all orders of application of functions yield the same output from an input, thus $F_i(F_j(x))=F_j(F_i(x))$, which is to say, the computation of one function does not depend on the results of the computation of another function. For numeric functions, $|1/x|=1/|x|$ so a system with "absolute value" and "multiplicative inverse" is time-invariant, but $succ(\sqrt{x})\neq\sqrt{succ(x)}$, thus "successor" and "square root" form a time-variant system. The notion of time invariance can be interpreted to refer to real time, but can also be viewed abstractly as referring to logical priority. N-stratality refers to the number of representations involved in computing a form. Most phonological theories are at least bi-stratal, having input and output representations, although DP seems not to have an input representation, thus would be monostratal.[29] Classical OT, Kimmo-style two-level phonology, and the DMH would appear to be bistratal, having just input and output representations. Two-level phonology might also be considered to have one representation with two aspects, the input and the output, thus Finnish *kammat* could be a single representation *kaMpat:kammat* where the substring to the right of the colon is what is pronounced. In like fashion, the Yawelmani OT candidate-chain <c'u:mhin, c'u:mhun, c'o:mhun, c'omhun> could be considered a single representation, only the last part of which is pronounced, but it is a representation with at least as many parts as a standard rule-based derivation. Without a clear definition of what constitutes a single representation, it is easy to achieve monostratality by conjoining derivational steps into one complex representational object.

Some constraint-based theories have a small fixed number of representations greater than two in the computation of an output, for example Harmonic Phonology which has three levels of representation. Polystratal theories can be subdivided into those with automatic and non-automatic strata: P&P and the URH are polystratal because their computations have multiple representations and the theories are time-variant (there is a correct vs. incorrect sequence of application of functions in the theory), but the sequence in which functions are applied is theoretically given automatically. Derivational phonology is generally non-automatically polystratal, so in a grammar with K rules, there are K representations, although there have been attempts to proscribe explicit ordering in certain cases, for example in Lexical Phonology to ascribe properties to the lexical vs. post-lexical modules. In OT, the selection of an optimal form is time-variant because the results of computing the winner from the sequence of marks {**,{*}} is not the same as computing the winner from the sequence of marks {*,{**}}.

While representational enrichment potentially translates multi-step derivations into multi-aspect representations, two- and three-level theories such as Kimmo Morphology and Harmonic Phonology seem to depend minimally on representational embellishment to eliminate steps in the production of outputs, and thus stand as the clearest alternatives to rule theory, with respect to the derivationality issue. One reason for concern over ordering mentioned by Goldsmith (1993b: 6) is the "100-step limitation," which refers to the fact that neural activity is not infinitely fast, so there may in principle be a maximal number of ordered steps in a derivation. This would be a concern for a theory aspiring to modeling an actual mental process, but not all theories have such aspirations.

5.3 Universality, Negativity

While universality was, historically speaking, seen as having a tight connection to constraints, no such connection is logically mandated. The connection between constraint (vs. rule) and universality has the dubious status of a question-begging presumption in generative grammar – *by definition*, constraints ought to be universals. In classical rule theory, the substance from which rules are constructed is drawn from a universal alphabet, combined into rules according to universal principles of rule construction, and sometimes with parameters which are universally available choices. Certain specific rules may be pre-supplied in a form, such as in atomic rules. At the same time, some theories assume that constraints are universally provided, but P&P phonology also allows language-specific constraints, and there seems to be little implication that constraints in a DP grammar are all universal. There is likewise little evidence that fixed-level constraint theories actually hold that constraints are pre-given by UG.

Universality is unlikely to be a valid argument for constraints over rules for two reasons. First, whether one uses rules or constraints, if one subscribes to the idea that there is some version of Universal Grammar, that entails a universal machinery and vocabulary, be it syntactic forms or lists, rules, or constraints. So if grammar is based on rules and UG states what the form of rules is, then *of course* there will be universals in the formulation of rules; equally, if grammar is based on constraints and UG states what the form of constraints is, then *of course* there will be universals in constraints formulation. Second, repeatedly observed phonological facts which defy reduction to a property of rule syntax – the fact that nasalization is vastly more common than denasalization, post-nasal voicing is vastly more common than post-nasal devoicing, and languages seem eager to give onsets to syllables and not so eager to get rid of onsets – may be at a probabilistic advantage from the perspective of sound change (see Hale and Reiss 2008), without reflecting on the grammatical faculty.

Similarly, the historical association between constraints and negative expressions vs. rules and positive expressions may be reinforced by the ordinary language association between constraint = negative command, vs. instruction = positive command, but without a well-justified theory of the form of constraints/rules and propositions that they depend on, any instruction to act one way when certain

conditions hold can almost trivially be translated into a prohibition against action any other way when those conditions do not hold. More interesting questions to ask would be: Are all/any features two-valued or monovalent? Should structural descriptions include disjunction as well as conjunction?

6 Conclusions

One firm conclusion that can be reached regarding rules vs. constraints as a model of phonology is that it is easy to be distracted by non-essential details of a particular theoretical package. The general ideas of rule-based and constraint-based grammar are sufficiently open-ended that neither can be per se reasonably judged superior to the other. A detailed and extensive comparison of a specific rule-based theory and a specific constraint-based theory could be productive, but is not the purpose of this chapter. Such a comparison must start from explicit metaphysical commitments – whether we are modeling sets of data presumed to already exist, or processes that generate complex data from simple parts; what facts are to be explained (low-level phonetic detail, neutralizing processes, lexically and morphologically governed processes); and whether the theory of phonology is held to account for the effect of grammar-independent factors of perception, production, and learning.

ACKNOWLEDGMENTS

This chapter was prepared with partial support from CASTL at the University of Tromsø. I would like to thank the editors, Mary Bradshaw, Mark Hale, Robert Levine, Detmar Meurers, Bruce Morén-Duollá, Andrew Nevins, Mary Paster, and Charles Reiss for discussion of various aspects of this chapter.

NOTES

1 The notion of empirical adequacy is straightforward: either a theory can handle the facts of language, or it cannot. The criteria for aesthetic adequacy are not well studied in linguistics: it corresponds to the difference between "explaining the facts" versus "merely grinding out the forms."

2 Natural Generative Phonology and Declarative Phonology share such a commitment to surface-true generalizations. See Hudson (1975), Hooper (1976) for basic NGP and the suppletive treatment of surface-opaque phonological alternations. In summarizing the essentials of Declarative Phonology, which strictly requires all statements to be surface true, Scobbie, Coleman, and Bird (1996: 703) claim "In particular, by arguing that would-be phonological transformations are in fact suppletive or phonetic, the one-level view of phonology is made tenable," and see pp. 694, 696 for hypothetical

examples of deletion and counter-feeding. Suppletive lexical listing is required in order to adhere to the "non-destructiveness" requirement of declarative theories: see Section 4.

3 Certain arguments for surface structure constraints in Shibatani (1973) were based on speaker-behavior facts which are outside grammar, and have no value in theory-selection phonology accounting for grammatical data patterns rather than mental states. McCarthy (2002: 10) rejects the Bromberger and Halle (1997) criticism of OT as entailing the impossibility of sorting an infinite set in finite time, on the grounds that such a consideration is external to competence, being properly part of a performance model in his opinion.

4 See Hale and Reiss (2008) for general discussion of these two perspectives.

5 However, transformational rules are (originally) taken to be non-Markovian in a limited way – see Chomsky (1956, 1957) – because reference to "NP" is a non-Markovian reference to the fact that a certain terminal substring such as "The little dog" derives from applying the rule NP → Det Adj N.

6 In a number of works, the term "rule" is also used to refer to what we would now identify as a well-formedness constraint, for example in parts of Stanley (1967).

7 There have been proposals to connect constraints to specific rules, for example the proposal of Sommerstein (1974) that rules may be "motivated by," and thus refer to, constraints, the parametric approach of Archangeli and Pulleyblank (1994), or Yip's (1988) use of the OCP as a force guiding rules.

8 *C* refers to category, for example, "noun," "noun phrase," "sentence."

9 Grammatical boundaries such as + and # had the formal status of matrices assigned the value [–segment,±word boundary,±formative boundary] – Chomsky and Halle (1968: 364 ff.).

10 The subscript on the braces indicates that the segments on the left become respectively the segments on the right, and that the changes are not random and unordered.

11 Whether braces themselves abbreviate conjunctively or disjunctively applied sub-rules is unclear. The Ordering Hypothesis (Chomsky and Halle 1968: 396) asserts that the sub-rules of a rule schema are applied disjunctively, but the following text fails to explicitly list braces as inducing disjunctivity. Chomsky (1965: 121) explicitly claims that braces abbreviate conjunctively-applied sub-rules.

12 It was a matter of some controversy whether the theory should allow general variables of the form X, or only abbreviatory expressions such as $(C_0V_0)_0$. See Odden (1977, 1980).

13 Since Howard (1972), questions have arisen as to the nature of the facts: see Archangeli and Pulleyblank (1994), Nevins (2004).

14 The caveat "phonologically-conditioned" also raises the possibility that phonological rules with non-phonological conditions may be subject to different principles.

15 This could be stated in standard notation as /VC [–voice]__ X, except that sonorants, which are voiced, also allow deletion of schwa.

16 This is not to say that all practitioners adhered to the principle that matrices should be fully specified: see for example Ringen (1975).

17 This rule must apply simultaneously: in an iterative model, the "free tone" condition would be invalidated after the first application of the rule. This is one of many cases of the critical interdependence of ideas in phonological theorizing, where the validity of one theory depends on the validity of an auxiliary proposition, which the competing theory does not depend on.

18 It should be noted that no argument has ever been given that there is categorical vowel deletion in *potato*, and the conclusion is based on the fact that in fast speech, there is

not usually any voicing on the vowel between voiceless stops. This is analogous to the problem of high-vowel vowel devoicing/deletion in Japanese, see Vance (1987).

19 This allows expression of the condition "is linked to something," where prior usage had required explicit specification of the two things linked.

20 Reiss proposes that both universal and existential quantifiers are required, to formulate the Non-Identity Condition – $\exists F_i \in G$ s.t $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$ – and the Identity Condition – $\forall F_i \in G$ s.t. $[(\alpha F_i)_1] = [(\beta F_i)_2]$. This formalism predicts two unattested conditions, Variable Partial Identity – $\exists F_i \in G$ s.t $[(\alpha F_i)_1] = [(\beta F_i)_2]$ where at least one feature must be the same – and Complete Non-Identity – $\forall F_i \in G$ s.t. $[(\alpha F_i)_1] \neq [(\beta F_i)_2]$ where all features must be non-identical. Reiss proposes a functional explanation for the nonexistence of the latter two classes. It is worth pointing out that this can also be formally explained. Exploiting DeMorgan's Laws, the Identity Condition can be equivalently expressed as $\neg \forall F_i \in G$ s.t. $[(\alpha F_i)_1] = [(\beta F_i)_2]$. Given that, Identity and Non-Identity are a single proposition $\forall F_i \in G$ s.t. $[(\alpha F_i)_1] = [(\beta F_i)_2]$ or its negation. If the formal theory only employs the notion of feature Identity, not Non-Identity, and only employs universal quantifiers, not existential quantifiers, then all and only the attested classes of identity conditions can be formalized.

21 It is not inconceivable that homorganic syncope could be reduced to "syncopate only if an OCP violation results," but that would be counter to the general trend on constraint-driven approaches that rules are triggered by constraints only if the rule increases harmony, not exacerbates constraint violation. The problem of refining the degree of identity remains, since identity effects variably ignore features which are structurally subordinate to the presumed shared node, such as voicing or retroflexion.

22 Rules A and B are complements iff the extension of the intersection of the structural descriptions is equivalent to the union of the extensions of the two structural descriptions.

23 Strictly speaking, the DMH probably cannot be falsified since in the worst case one could simply list all of the input-output pairs of a language; but it can be shown that the rule system entailed by the theory is intolerable, in that it misses major generalizations. A test case might be possible which involved the phrasal phonology of actually unbounded clauses, of the type discussed in Odden (2000): such rules are very rare.

24 Karttunen has no discussion of the remainder of the conditioning environment, viz. "when in the onset of a closed syllable."

25 These are seen as co-existing levels of representation.

26 This is modeled in Goldsmith's approach through a connectionist-type equation involving inherent activations and lateral inhibitions.

27 See Port and O'Dell (1985) for arguments that some claimed neutralizing rules of phonology are not actually neutralizing. See Liphola (2001) for experimental evidence confirming that Makonde vowel reduction is acoustically and perceptually neutralizing.

28 With respect to expressing a single rule in terms of a set of constraints: a derivational grammar includes not just rules, but also ordering statements, which are not the topic of this chapter.

29 Whether this is actually so depends on how "single representation" is defined – see below – and for DP, how neutralizing processes are formally handled.

2 Opacity and Ordering

ERIC BAKOVIĆ

1 Introduction

Few notions in phonological theory have received as much attention in the literature as OPACITY. In the almost 40 years since Kiparsky (1971, 1976) offered the definition given in (1), the bulk of the attention paid to opacity has been relatively recent and has been fueled by the field's massive (but incomplete) shift from the rule-based serialism framework of *The Sound Pattern of English* (Chomsky and Halle 1968) to the constraint-based parallelism framework of Optimality Theory (Prince and Smolensky 1993).

(1) Opacity (Kiparsky 1976: 79)

A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \ / \ C__D$ is OPAQUE if there are surface structures with either of the following characteristics:
a. instances of A in the environment $C__D$.
b. instances of B derived by $\mathbb{P}$ that occur in environments other than $C__D$.

According to (1), the opacity of a (hypothesized) rule $\mathbb{P}$ can be formally diagnosed by comparing the set of (predicted) surface representations with the generalization expressed by $\mathbb{P}$: to say that $\mathbb{P}$ is opaque is to say that the applicability or application of $\mathbb{P}$ is (somehow) obscured on the surface. Kiparsky's substantive claim was that an opaque rule $\mathbb{P}$ is *difficult to learn*, either (1a) because there are surface counterexamples to $\mathbb{P}$'s applicability, or (1b) because there are surface contexts in which $\mathbb{P}$'s application is not motivated.

Kiparsky's support for this substantive learnability claim was a set of examples of language change in which previously opaque rules become transparent. More

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

specifically, Kiparsky identified two pairwise orders between rules made possible by the rule-based serialism framework, COUNTERFEEDING and COUNTERBLEEDING, and argued that each order (i) results in a particular type of opacity ((1a) and (1b), respectively), and (ii) tends to change over time to the corresponding reverse, transparent order (FEEDING and BLEEDING, respectively). These orders (and Kiparsky's claims) are discussed more extensively in Section 2.

If there's only one thing that phonologists have learned from Kiparsky's work on the subject of opacity, it is to equate opacity of type (1a) with counterfeeding and opacity of type (1b) with counterbleeding. My aim here is to demonstrate that these equations are falsified in both directions: in Section 3 I show that not all cases of type (1a) opacity result from counterfeeding and that not all cases of counterfeeding result in opacity of either type, and in Section 4 I show that not all cases of type (1b) opacity result from counterbleeding and that not all cases of counterbleeding result in opacity of either type. This demonstration reveals a very different, more complex, and more complete picture of what opacity is than previously conceived. This is a significant result because opacity's original raison d'être is Kiparsky's claim that an opaque rule is difficult to learn. This claim is meaningful and testable only insofar as we have a clear understanding of what is and what is not an instance of an opaque rule, and what an account of such an instance, in turn, should look like.

2 Pairwise Rule Ordering

The central principle of rule-based serialism is rule ordering. Bromberger and Halle's (1989: 58–59) informal definition of rule ordering, given in (2), suffices for our purposes.

(2) Rule ordering (Bromberger and Halle 1989: 58–59)

> Phonological rules are ordered with respect to one another. A phonological rule R does not apply necessarily to the underlying representation; rather, R applies to the derived representation that results from the application of each applicable rule preceding R in the order of the rules.

There are four recognized non-trivial pairwise ordered rule relations in rule-based serialism: feeding, bleeding, counterfeeding, and counterbleeding. These are defined informally in (3).[1]

(3) Pairwise ordered rule relations (adapted from McCarthy 2007b)

> Given two rules $\mathbb{A}$, $\mathbb{B}$ such that $\mathbb{A}$ precedes $\mathbb{B}$,
> a. $\mathbb{A}$ FEEDS $\mathbb{B}$ iff $\mathbb{A}$ creates additional inputs to $\mathbb{B}$.
> b. $\mathbb{A}$ BLEEDS $\mathbb{B}$ iff $\mathbb{A}$ eliminates potential inputs to $\mathbb{B}$.
> c. $\mathbb{B}$ COUNTERFEEDS $\mathbb{A}$ iff $\mathbb{B}$ creates additional inputs to $\mathbb{A}$.
> d. $\mathbb{B}$ COUNTERBLEEDS $\mathbb{A}$ iff $\mathbb{B}$ eliminates potential inputs to $\mathbb{A}$.

Note that counterfeeding and counterbleeding are *counterfactual inverses* of feeding and bleeding, respectively, because counterfeeding *would be* feeding and counterbleeding *would be* bleeding if the two rules involved were ordered in the opposite way. The terminology, though notoriously difficult to learn, is thus not completely misleading.

Two rules may interact in different ways in different derivations. Consider (4), for example. In (4a), Deletion feeds Palatalization: deletion of the /u/ crucially places the preceding /t/ before a [–back] vowel. In (4b), on the other hand, Deletion bleeds Palatalization: the deleted /i/ is [–back] and thus would have induced palatalization of the preceding /t/ if it hadn't been deleted. In both (4c) and (4d), the two rules are mutually non-affecting: in (4c), neither vowel is [–back] and so the /t/ is never in a context to be palatalized; in (4d), both vowels are [–back] and so the /t/ is in a context to be palatalized either way.

(4) Feeding and bleeding in different derivations (hypothetical)

		a. /tue/	b. /tio/	c. /tou/	d. /tei/
Deletion:	V → Ø / __ V	Ø	Ø	Ø	Ø
Palatalization:	t → ʧ / __ [–back]	ʧ			ʧ
		[ʧe]	[to]	[tu]	[ʧi]

Reversing the order of these two rules, as in (5), we get counterfeeding and counterbleeding in different derivations.[2] In (5a), Deletion counterfeeds Palatalization: deletion of the /u/ places the preceding /t/ before a [–back] vowel, but too late for Palatalization to do anything about it. In (5b), on the other hand, Deletion counterbleeds Palatalization: the deleted /i/ is [–back] and thus induces palatalization of the preceding /t/ before deleting. In both (5c) and (5d), the two rules are again mutually non-affecting, just as in (4) above.

(5) Counterfeeding and counterbleeding in different derivations (hypothetical)

		a. /tue/	b. /tio/	c. /tou/	d. /tei/
Palatalization:	t → ʧ / __ [–back]		ʧ		ʧ
Deletion:	V → Ø / __ V	Ø	Ø	Ø	Ø
		[te]	[ʧo]	[tu]	[ʧi]

Although (3) constitutes a useful picture of the typology of possible ordered rule relations predicted by the central principle of rule-based serialism in (2), it is still defined (almost) exclusively in terms of interactions between just two ordered rules. I hardly hesitate to qualify this statement because most if not all definitions of pairwise ordered rule relations provided in textbooks and in the scholarly literature are insufficiently precise about situations involving more than two rules, which may counterintuitively fit or not fit a given definition.[3] But the fact remains that the bulk of the relevant literature focuses on pairwise interactions.

There have been two significant proposals for classifying the ordering relations in (3). The first was the relatively formal hypothesis that "rules tend to shift into the order which allows their fullest utilization in the grammar" (Kiparsky 1968c:

200). This privileges feeding and counterbleeding orders, grouping them together as "unmarked" because these are the orders in which both rules apply non-vacuously – that is, in which the two rules are both utilized, as can be appreciated from the feeding derivation in (4a) and the counterbleeding derivation in (5b) above. Conversely, bleeding and counterfeeding orders are "marked" because these are the orders in which one of the two rules fails to apply non-vacuously, as can be appreciated from the bleeding derivation in (4b) and the counterfeeding derivation in (5a).

There were several challenges to Kiparsky's "maximal utilization" hypothesis; see Ken stowicz and Kisseberth (1977: 159ff.) for an informative summary critique. Kiparsky's response was a relatively substantive second hypothesis, that "rules tend to be ordered so as to become maximally transparent" (Kiparsky 1971: 623). A transparent rule is one that does not meet either of the two conditions defined in (1) above, repeated in (6) below.

(6) Opacity, repeated from (1)

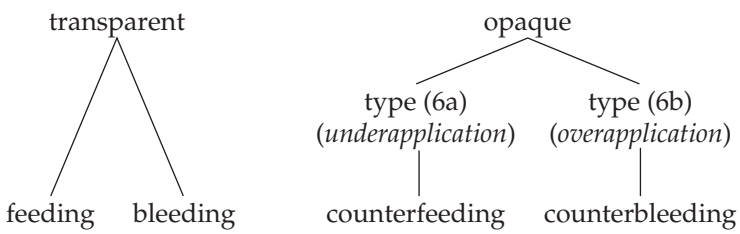
A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \ / \ C__D$ is OPAQUE if there are surface structures with any of the following characteristics:

a. instances of A in the environment $C__D$.
b. instances of B derived by $\mathbb{P}$ that occur in environments other than $C__D$.

Kiparsky hypothesized that diachronic change proceeds from harder-to-learn opacity-promoting rule orders to easier-to-learn transparency-promoting ones, modulo potentially conflicting principles such as paradigm uniformity. Kaye (1974, 1975), Kisseberth (1976), and Kenstowicz and Kisseberth (1977: 170ff.) question the overall learnability claim by pointing out that phonological opacity often helps to maintain lexical contrasts (which one might think of as "semantic transparency"); see Łubowicz (2003a) for a recent rearticulation of this view.[4]

McCarthy (1999) adapts a couple of terms from work on reduplication by Wilbur (1973), UNDERAPPLICATION and OVERAPPLICATION, to elucidate the two types of opacity in (6).[5] Type (6a) describes situations in which there are surface representations to which $\mathbb{P}$ could apply non-vacuously; $\mathbb{P}$ has thus *underapplied*. Type (6b) describes situations in which there are surface representations to which $\mathbb{P}$ has applied non-vacuously, but which do not otherwise meet $\mathbb{P}$'s structural description; $\mathbb{P}$ has thus *overapplied*. Kiparsky's explicit and subsequently generally accepted classification of the four pairwise rule interactions in (3) is shown in (7).

(7) Classification of pairwise ordered rule interactions (Kiparsky 1971, 1976)



In the next two sections I demonstrate that the classification of pairwise ordered rule interactions, in (7) is misleading at best. Counterfeeding is but one of several devices that can be and have been used to describe actual examples meeting the definition of underapplication in (6a), and counterfeeding does not always lead to underapplication (Section 3). Similarly, counterbleeding is not the only way to describe actual examples meeting the definition of overapplication in (6b), and counterbleeding does not always lead to overapplication (Section 4).

3 Underapplication and Counterfeeding

The definitions of underapplication opacity in (6a) and of the counterfeeding relation in (3c) are repeated (in suitably modified forms) in (8) and (9), respectively.

(8) A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \ / \ C__D$ UNDERAPPLIES if there are surface structures with instances of A in the environment $C__D$.

(9) $\mathbb{B}$ COUNTERFEEDS $\mathbb{A}$ iff $\mathbb{B}$ creates additional inputs to $\mathbb{A}$ and $\mathbb{A}$ precedes $\mathbb{B}$.

I begin in Section 3.1 by explaining how some examples of counterfeeding as defined in (9) result in underapplication as defined in (8). Then I demonstrate that counterfeeding is not the only source of underapplication. In Section 3.2 I discuss various types of BLOCKING, the most obvious type of underapplication that is not typically categorized as such in the literature, and in Section 3.3 I discuss a handful of other phenomena that also arguably contribute to underapplication opacity: the restriction of a rule to particular lexical classes or levels, rule exceptions, and rule optionality. Finally, I demonstrate in Section 3.4 that counterfeeding does not always lead to underapplication opacity, at least not as underapplication is defined in (8).

3.1 Counterfeeding

The counterfeeding relation in (9) describes situations where a later-ordered rule $\mathbb{B}$ creates representations to which an earlier-ordered rule $\mathbb{A}$ could have applied non-vacuously; modulo the action of other, even later rules (see Section 3.4), $\mathbb{A}$ underapplies in such situations. This was exemplified by the hypothetical derivation of /tue/ in (5a) above: Deletion creates an additional input to Palatalization, but because Palatalization precedes Deletion the result is a surface structure, [te], with a voiceless coronal stop before a front vowel – the structural description of Palatalization. Palatalization has thus underapplied in this derivation.

Following McCarthy (1999), I distinguish COUNTERFEEDING ON ENVIRONMENT from COUNTERFEEDING ON FOCUS interactions (see also Baković 2007: 221ff.). In a rule of the form $A \rightarrow B \ / \ C__D$, the *focus* is A, the element to be changed by the rule, and the *environment* is $C__D$, the necessary context surrounding the focus. In counterfeeding on environment interactions the later-ordered rule $\mathbb{B}$ creates

the environment of the earlier-ordered rule 𝔸, and in counterfeeding on focus interactions 𝔹 creates the focus of 𝔸. The main significance of this distinction is that cases of counterfeeding on focus have comparably successful accounts without ordering, as will be briefly noted in Section 3.1.2.

3.1.1 Counterfeeding on Environment Consider as an example of counterfeeding on environment the following two rules of Lomongo.

(10) Counterfeeding in Lomongo (Hulstaert 1961; Kenstowicz and Kisseberth 1979)

	a. /o+bina/	b. /o+isa/	c. /ba+bina/
Gliding: [–low] → [–syll] / __V		w	
Deletion: $\begin{bmatrix}+\text{voi}\\-\text{son}\end{bmatrix}$ → ∅ / V __	∅		∅
	[o+ina]	[w+isa]	[ba+ina]

Glosses: (10a) 'you (sg.) dance', (10b) 'you (sg.) hide', (10c) 'they dance'

The derivations in (10b–c) illustrate the independent action of each of the rules: gliding applies alone in (10b) and Deletion applies alone in (10c), with no interaction in either case. In (10a), Deletion counterfeeds Gliding by creating the environment (a following vowel) that Gliding could have used to apply to the /o/. Gliding thus underapplies because there are surface representations with non-low prevocalic vowels that have not become glides.

There are also more complex interactions involving counterfeeding on environment, for instance where 𝔸 feeds 𝔹 but 𝔹 in turn counterfeeds 𝔸. I borrow from Kavitskaya and Staroverov (2010) the term "fed counterfeeding" to refer to this type of interaction. An example of FED COUNTERFEEDING ON ENVIRONMENT is found in Lardil, as shown in (11).

(11) Fed counterfeeding in Lardil (Hale 1973; Kavitskaya and Staroverov 2010)[6]

	a. /dibirdibi/	b. /yiliyili/	c. /wangalk/
Apocope: V → ∅ / σ σ __ #	∅	∅	
Deletion: [–apical] → ∅ / __ #	∅		∅
	[dibirdi]	[yiliyil]	[wangal]

Glosses: (11a) 'rock cod', (11b) 'oyster species', (11c) 'boomerang'

The derivations in (11b–c) again illustrate the independent action of each of the rules: in (11b), application of Apocope leaves a word-final apical consonant behind, which is not subject to Deletion; in (11c), there is no word-final vowel before or after application of Deletion. In (11a), Apocope feeds Deletion: removal of the word-final vowel places the preceding non-apical consonant in a position to be deleted. But Deletion also counterfeeds Apocope here: deletion of the non-apical consonant places the preceding vowel in a position to also be removed by Apocope, but Apocope does not apply to this vowel. Apocope thus underapplies because there are surface representations with word-final vowels.

3.1.2 Counterfeeding on Focus Now consider as an example of counterfeeding on focus the following rules of Western Basque.

(12) Counterfeeding in Western Basque (de Rijk 1970; Hualde 1991; Kawahara 2002)

		a. /alaba+a/	b. /seme+e/
Raising-to-High:	[–low] → [+high] / __V		i
Raising-to-Mid:	[+low] → [–low] / __V	e	
		[alabe+a]	[semi+e]

Glosses: (12a) 'daughter', (12b) 'son'

The derivation in (12b) illustrates the independent action of Raising-to-High, which applies alone here to raise the prevocalic mid vowel. In (12a), Raising-to-Mid applies and counterfeeds Raising-to-High by changing the focus to a mid vowel that Raising-to-High could have applied to if it were later in the order. Raising-to-High thus underapplies because there are surface representations with mid prevocalic vowels that have not become high.

Examples of counterfeeding on focus like this one, particularly when the environments of the rules are the same, are referred to as CHAIN SHIFTS: underlying *A* becomes *B* and underlying *B* becomes *C*, but an *A* that becomes a *B* does not go on to become a *C*.

A comparably successful alternative to the ordering analysis of chain shifts recognizes the scalar nature of the dimensions along which chain shifts tend to occur (Kirchner 1996; Baković 1996; Gnanadesikan 1997; Kawahara 2002; Moreton and Smolensky 2002): movement toward the target end of the scale, even if it is not all the way, is better than no movement at all. In Western Basque, for example, the relevant scale is that of vowel height and the target end of the scale is a high vowel; both underlying mid and underlying low vowels aim in the right direction, though only mid vowels manage to hit the target.

Another comparably successful alternative capitalizes on the fact that chain shifts are *contrast-preserving* (Łubowicz 2003a, b): the fact that underlying *A* surfaces as *B* and underlying *B* surfaces as *C* means that the underlying contrast between *A* and *B* is manifested as a contrast, albeit a shifted one, on the surface. (See Łubowicz-Baković 2011, and references therein for more details on chain shifts and their analysis.)

There are also examples of FED COUNTERFEEDING ON FOCUS, for example in Nootka:

(13) Fed counterfeeding in Nootka (Sapir and Swadesh 1978; McCarthy 1999, 2003, 2007b)

	a. /muː q/	b. /ħaju+q i/	c. /ɬaː kʷ+ʃit͡ɬ/
Labialization:			
[+dors] → [+rnd] / [+rnd] __	q^w	q^w	
Delabialization:			
[+dors] → [–rnd] / __]$_\sigma$	q		k
	[muː q]	[ħaju+q^wi]	[ɬaːk +ʃit͡ɬ]

Glosses: (13a) 'throwing off sparks', (13b) 'ten on top', (13c) 'to take pity on'

The derivations in (13b–c) yet again illustrate the independent action of each of the rules. In (13b), Labialization applies to a dorsal that is preceded by a round vowel but is not syllable-final, and so Delabialization is inapplicable; in (13c), Delabialization applies to a syllable-final dorsal that is not preceded by a round vowel, and so Labialization is inapplicable. In (13a), Labialization both feeds and is counterfed by Delabialization: the dorsal is preceded by a round vowel and so it labializes, but this creates a syllable-final labialized dorsal that is subsequently delabialized – which puts the dorsal back in the position of being non-vacuously subject to Labialization. Thus, even though Labialization "applies" in the sense that it makes a non-vacuous change during the course of the derivation, this rule underapplies in the specific sense defined in (8). (See Section 3.4 for discussion of an example of the converse situation: a rule that does *not* make a non-vacuous change during the course of the derivation but that still does not underapply in the sense of (8).)

Examples of fed counterfeeding on focus are more commonly referred to as DUKE OF YORK DERIVATIONS (Pullum 1976; McCarthy 2003): an underlying *A* becomes *B* only to end up as *A* again. As with chain shifts, there is a comparably successful alternative to the ordering analysis of Duke of York derivations, involving the conflict-adjudication mechanism of constraint ranking in Optimality Theory (McCarthy 1999, 2003a, 2007b): in Nootka, for example, the markedness constraint driving Delabialization must be ranked higher than the markedness constraint driving Labialization. (A particular subset of Duke of York derivations is also amenable to disjunctive blocking analysis; see Section 3.2.1 below.)

3.2 Blocking

Cases of counterfeeding like those discussed above have convinced many phonologists that underapplication opacity is fully accounted for by rule ordering; after all, if a demonstrably active rule's input structural description is met by a surface representation, it makes sense to think that another, later-ordered rule created that representation. But there are also sources of underapplication other than counterfeeding, all of which have received ample attention in the phonological literature. I begin with the most obvious such source, blocking.

The very definition of blocking belies its contribution to underapplication: a rule is said to be blocked when it fails – by some principle or mechanism – to apply to a form that meets its input structural description; thus, a derivation in which a given rule $\mathbb{P}$ has been blocked may result in a surface representation to which $\mathbb{P}$ underapplies. (I say "*may* result" because another, later-ordered rule could rid the surface of representations that meet the structural description of $\mathbb{P}$. Counterfeeding can be made transparent in this way; see Section 3.4.)

I discuss here three types of blocking. The first is DISJUNCTIVE BLOCKING (Section 3.2.1), in which a rule is blocked if a strictly more specific conflicting rule is also applicable. The second is NON-DERIVED ENVIRONMENT BLOCKING (Section 3.2.2), in which a rule is blocked if its structural description is not derived phonologically or morphologically. The third is (for lack of a better term)

DO-SOMETHING-EXCEPT-WHEN BLOCKING (Section 3.2.3), in which a rule is blocked from creating structures that for independent reasons are not allowed to surface. (The closely-related phenomenon of DO-SOMETHING-ONLY-WHEN TRIGGERING also involves underapplication, as also noted in Section 3.2.3.) Each of these well-established phenomena has required the postulation of principles or mechanisms beyond rule ordering to account for it; given that each type of blocking (and triggering) contributes to underapplication, then, it is clear that rule ordering is insufficient to account for all cases of opacity.

3.2.1 Disjunctive Blocking Disjunctive blocking has a long and celebrated history in phonological theory (see Baković, forthcoming, for detailed discussion). It all started with the analysis of stress in Chomsky, Halle, and Lukoff (1956), Chomsky and Halle (1968), and Halle and Keyser (1971). Consider the Latin stress rules in (14), stated in standard *SPE* notation (after Anderson 1974: 97).

(14) Latin stress rules

a.	$V \rightarrow [+\text{stress}] \; / \; __ \; C_0\check{V}C_0^1VC_0\#$	(stress the antepenult if the penult is light)
b.	$V \rightarrow [+\text{stress}] \; / \; __ \; C_0VC_0\#$	(stress the penult)
c.	$V \rightarrow [+\text{stress}] \; / \; __ \; C_0\#$	(stress the ultima)

Any form fitting the structural description of one of the longer rules in (14) also fits the structural description of any shorter rule. Application of these rules to any form that meets the structural description of more than one of the rules will thus result in multiple stresses on the form, regardless of the order of the rules. However, only (14a) applies to words that fit the structural descriptions of all three rules (*pa'tricia, 'reficit*), only (14b) applies to words that fit the structural descriptions of (14b,c) but not that of (14a) (*re'fectus, re'fēcit, 'aqua, 'amō*), and only (14c) applies to words that fit its structural description and not those of the other two rules (*'mens, 'cor, 'rē*). Application of a shorter, more general rule must thus be blocked by application of a longer, more specific rule; the shorter, more general rules thus underapply, again in a way that cannot be accounted for with rule ordering alone.

Other types of examples of disjunctive blocking were identified by Anderson (1969, 1974) and Kiparsky (1973), and all such cases have since been generally accounted for by (some version of) Kiparsky's ELSEWHERE CONDITION (Kiparsky 1973, 1982a). (Complementary stress rules such as those in (14), on the other hand, were eventually superseded by the interaction of principles of metrical phonology, as noted by Kiparsky 1982a: 173, footnote 2.) The Elsewhere Condition imposes disjunctive ordering between two rules the structural changes of which are incompatible and the structural descriptions of which are in a proper inclusion relationship. Many, but not all, such examples can in fact also be accounted for by a Duke of York derivation (recall Section 3.1.2). An example of this kind is the interaction between Trisyllabic Shortening and *CiV*-Lengthening in English (Chomsky and Halle 1968; Kenstowicz 1994a).

(15) English rules (adapted from Kenstowicz 1994a: 218)

a. Trisyllabic Shortening — e.g. *o*('*pāque*) ~ *o*('*păci*)*ty*

$$V \rightarrow \breve{V} \;/\; \underset{(\text{'}\sigma}{\underset{|}{__}}\ C_0\ \underset{\sigma)}{V}$$

b. *CiV*-Lengthening — e.g. (*'remĕ*)*dy* ~ *re*('*mēdi*)*al*

$$\begin{bmatrix} V \\ -\text{high} \end{bmatrix} \rightarrow \bar{V} \;/\; \underset{(\text{'}\sigma}{\underset{|}{__}}\ C\ i\ \underset{\sigma)}{V}$$

Application of these rules to forms that meet both structural descriptions results in the right surface representations, whether the rules are ordered normally (= conjunctively) or disjunctively. I explain this fact in what follows, employing as key examples the forms /*remĕdi+al*/ and /*jōvial*/ (→ |*re*('*mĕdi*)+*al*| and |(*'jōvi*)*al*| after footing, respectively).

Kenstowicz (1994a: 218) advocates a disjunctive analysis, mediated by the Elsewhere Condition. The structural changes of the rules are incompatible: one rule shortens vowels while the other lengthens them. Moreover, the structural description of (*CiV*-)Lengthening is properly included in that of (Trisyllabic) Shortening: both apply to the heads of bisyllabic feet, but Lengthening applies more specifically to a [–high] head of a foot the non-head of which is an /*i*/ in hiatus. Lengthening thus blocks Shortening, and Lengthening therefore applies alone to |*re*('*mĕdi*)+*al*| (→ [*re*('*mēdi*)*al*]) and |(*'jōvi*)*al*| (→ [(*'jōvi*)*al*]).

Chomsky and Halle (1968: 181, 240ff.) propose a conjunctive analysis, with extrinsic ordering between the two rules.[7] Shortening applies first and gives the intermediate representations |*re*(*mĕdi*)+*al*| and |(*jŏvi*)*al*|; Lengthening then undoes the effects of Shortening in these cases, rendering the correct surface representations [*re*(*mēdi*)*al*] and [(*jōvi*)*al*]. This is a clear example of fed counterfeeding on focus (recall Nootka, Section 3.1.2, (13)): Lengthening feeds Shortening which in turn counterfeeds Lengthening, which thus underapplies.

There are other examples of disjunctive blocking that can be shoe-horned into conjunctive analyses, but only at the expense of the descriptive adequacy of the individual rules themselves. Consider, for example, the interaction between Assimilation and Deletion in Diola Fogny (Sapir 1965, Kiparsky 1973), starting with the disjunctive analysis in (16).

(16) Diola Fogny rules (disjunctive analysis, adapted from Kiparsky 1973: 98)

a. Assimilation — e.g. /ni+gam+gam/ → [nigaŋgam] 'I judge'

$$\begin{bmatrix} C \\ +\text{nasal} \end{bmatrix} \rightarrow [\alpha\text{place}] \;/\; __ \begin{bmatrix} -\text{cont} \\ \alpha\text{place} \end{bmatrix}$$

b. Deletion — e.g. /let+ku+ʤaw/ → [lekuʤaw] 'they won't go'

$$C \rightarrow \emptyset \;/\; __\ C$$

The structural description of Assimilation is properly included in that of Deletion: both apply to preconsonantal consonants, but Assimilation applies more specifically to nasals followed by non-continuants. Moreover, the structural changes of the two rules are incompatible: a consonant can either be assimilated or deleted, but not both (not discernibly, anyway). Assimilation thus applies alone when applicable, blocking Deletion.

Unlike the English rules in (15), the Diola Fogny rules as stated in (16) cannot be ordered conjunctively: under either order, Deletion will delete all preconsonantal consonants, whether or not they (were destined to) undergo Assimilation. A conjunctive analysis of the interaction between these two rules requires rules as stated and as ordered in (17).

(17) Diola Fogny rules (conjunctive analysis, adapted from Kiparsky 1973: 97)

a. Deletion′

$$\begin{bmatrix} \text{C} \\ \text{<+nasal>} \end{bmatrix} \rightarrow \emptyset \; / \; __ \begin{bmatrix} \text{C} \\ \text{<+cont>} \end{bmatrix}$$

e.g. /na+laɲ+laɲ/ → [nalalaɲ] ‘he returned’

b. Assimilation′

$$\text{C} \rightarrow [\alpha\text{place}] \; / \; __ \begin{bmatrix} \text{C} \\ \alpha\text{place} \end{bmatrix}$$

e.g. /ku+bɔn+bɔn/ → [kubɔmbɔn] ‘they sent’

Deletion′ deletes a nasal only if it is followed by a continuant, and otherwise deletes all preconsonantal consonants. The relevant residue of this rule – nasals followed by non-continuants – is then passed on conjunctively to Assimilation′. This means that Assimilation′ need not specify the non-continuancy of the consonant being assimilated to, because Deletion′ will have already removed the relevant strings from consideration. The continuancy of the following consonant is thus a condition on Deletion′ under this conjunctive analysis, as opposed to being a condition on Assimilation as it is in the disjunctive analysis – and herein lies the problem with the conjunctive analysis. That the following consonant must be [–cont] in order for Assimilation to apply in (16a) is a natural condition on nasal place assimilation rules (Padgett 1994), but the condition on Deletion′ in (17a) – that the following consonant should be [+cont] if the consonant-to-be-deleted is [+nasal] – is not similarly justified.

In summary, disjunctive blocking represents yet another example of underapplication that cannot be accounted for with rule ordering alone. Even factoring out examples like the Latin case in (14), instead accounting for them via the interaction of principles of metrical phonology, and examples like the English case in (15), which can be inconsequentially reanalyzed as a Duke of York derivation, there remains a residue of examples like the Diola Fogny case in (16) that are best described as involving the underapplication of a rule due to disjunctive blocking by another, rather than conjunctive ordering with respect to another.

3.2.2 Non-derived Environment Blocking A classic example of non-derived environment blocking is found in Finnish (Kiparsky 1976, 1993) and is shown in (18).

(18) Non-derived environment blocking in Finnish

		a. /tilat+i/	b. /äiti/	c. /vete/
Raising:	e → i / __ #			i
Assibilation:	t → s / __ i	≀ s	≀	s
		[tilas +i]	[äiti]	[vesi]

Glosses: (18a) 'ordered', (18b) 'mother', (18c) 'water'

The examples in (18) show that Assibilation only applies if its structural description is morphologically or phonologically derived; that is, only when the conditions for (non-vacuous) application of the rule are met by virtue of the concatenation of morphemes, as in (18a), or by the application of a prior phonological rule, as in (18c). The morpheme-final /t/ in (18a) assibilates because the conditioning vowel is in a separate morpheme; the initial /t/ does not assibilate, however – as indicated by the ad hoc '≀' symbol – because the would-be conditioning vowel is in the same morpheme. The example in (18b) has a /t/ in virtually the same phonological context as the assibilated /t/ in (18a) and yet it does not assibilate because, like the unassibilated initial /t/ of (18a), the conditioning vowel is in the same morpheme. Finally, the /t/ in (18b) assibilates because the conditioning vowel is derived by the earlier application of Raising. Assibilation clearly underapplies in Finnish, given that there are surface representations that could have undergone Assibilation but have not.

Note that the conditions that hold of non-derived environment blocking are essentially the opposite of those that hold of counterfeeding. In cases of counterfeeding, earlier-derived strings undergo a rule that later-derived strings do not; ordering this rule earlier than another rule that is responsible for those later-derived strings is thus possible. In cases of non-derived environment blocking, by contrast, later-derived strings (whether by morpheme concatenation or by phonological rule) undergo a rule that earlier-derived strings do not. Rule ordering is clearly insufficient to the task in this case: early ordering can only hope to achieve counterfeeding-type underapplication, and late ordering will if anything only increase the set of forms to which the relevant rule can apply. As the ample literature on the topic attests, some additional principle ensuring the blocking of relevant rules in non-derived environments (or, alternatively, their application only in derived environments) is necessary within rule-based serialism, in the form of either the Revised Alternation Condition (Kiparsky 1976), the Strict Cycle Condition (Kean 1974; Mascaró 1976), a combination of lexical identity rules and the Elsewhere Condition (Kiparsky 1982), or the judicious use of underspecification and feature-filling rule application (Kiparsky 1993; cf. Poser 1993). (See Burzio-Baković 2011, and references cited there for more on non-derived environment blocking and its analysis.)

3.2.3 Do-Something-Except-When Blocking Do-something-except-when blocking encompasses a wide range of cases in which a rule is blocked from creating certain structures for independently-motivated reasons. It is usually motivated by the *general absence* of a particular structure in a language, one that is otherwise expected to be created by the rule in question. It differs from disjunctive blocking in that another rule (formally related or otherwise) is generally not involved, and it differs from non-derived environment blocking in that the relevant structures are generally blocked from being created across the board, not only in non-derived environments. But it is like both of these other forms of blocking in that it involves underapplication of the blocked rule.

The earliest argument for do-something-except-when blocking was made by Kisseberth (1970). In Yawelmani Yokuts (Newman 1944; Kuroda 1967; Kisseberth 1969), short vowels are deleted between consonants *except when* such deletion would result in a tautosyllabic consonant cluster (#CC, CCC, or CC#). One way to achieve this result is, of course, to build the blocking condition into the statement of the vowel deletion rule, the environment of which can be stated as VC__CV (a "doubly open syllable"), thereby including all but those contexts in which a tautosyllabic consonant cluster is in danger of being created. Kisseberth (1970) argues that this solution misses a significant generalization uniting a suite of rules in Yawelmani phonology that are either blocked or triggered (on which see below) by the avoidance of tautosyllabic consonant clusters. He argues instead that the environment of vowel deletion could instead be simplified to C__C, with the surrounding vowels of the more complex VC__CV environment being derivative properties of a CONSPIRACY.[8] To the extent that such derivative properties can indeed be factored out of the formal statement of the environment of a conspiracy-blocked rule, then, that rule underapplies.[9]

This is also true of rules that are blocked for other do-something-except-when reasons. For example, assimilation rules are often subject to the same conditions as the underlying segment inventory itself, such that the product of assimilation cannot be a segment outside the inventory. Vowel harmony rules offer some of the most consistent evidence for this. In the vowel inventory of the Fante variety of Akan (Stewart 1967, Clements 1981, O'Keefe 2003), all vowels have a [±ATR] pair /i ~ ɪ, e ~ ɛ, u ~ ʊ, o ~ ɔ/ except the low, [–ATR] vowel /a/. As a result, the [±ATR] vowel harmony rule is blocked from applying to /a/. In this case, this blocking condition can be built in to the statement of the focus of the vowel harmony rule by stipulating that it only applies to [–low] vowels, but this has been argued since at least Kiparsky (1981) to miss a significant generalization about the relationship between conditions on harmony and conditions on the inventory. Kenstowicz and Kisseberth (1977) discussed cases like this under the rubric of the DUPLICATION PROBLEM, explaining that, as with conspiracies, the rule-based serialism model of the time was forced to view this kind of relationship as a coincidence; later work addressed the duplication problem with the STRUCTURE PRESERVATION principle (Kiparsky 1981, 1982a, 1985).

A recently proposed subclass of do-something-except-when blocking is represented by what McCarthy (2003a) calls a GRANDFATHER EFFECT, whereby a rule is

blocked from creating a representation that is otherwise allowed to surface if specified underlyingly; these underlying forms are thus "grandfathered in." McCarthy uses voicing assimilation in Mekkan Arabic as an example (Abu-Mansour 1996, Bakalla 1973): underlying voiced obstruents assimilate to following voiceless ones (/ʔagsam/ → [ʔaksam] 'he swore an oath') but not vice-versa (/ʔakbar/ → [ʔakbar], *[ʔagbar] 'older'), even though voiced obstruents emerge unscathed if specified underlyingly (/ʔibnu/ → [ʔibnu] 'his son'). Grandfather effects are not independently motivated by conspiracies or inventory conditions, but McCarthy argues that they are motivated by universal markedness considerations: in the Mekkan Arabic case, the fact that voiced obstruents are marked by comparison with voiceless ones. To the extent that such markedness constraints can, like independently-motivated inventory conditions, be factored out of the rules that they block, then, rules of this kind also underapply.

Rules that are *triggered* by conspiracies or by inventory conditions (= "do something *only* when") also underapply, at least to the extent that the relevant derivative properties can be factored out of the formal statement of the environment of the triggered rule. For example, one of the conspiracy-triggered rules of Yawelmani discussed by Kisseberth (1970) epenthesizes a vowel after the first consonant of what would otherwise be a tautosyllabic consonant cluster; if the environment of the rule could thereby be reduced just to the position of epenthesis (to the effect that "in a sequence of one or more consonants, epenthesize after the first consonant"), then it would technically underapply in all sequences of one or more consonants that are not in danger of surfacing as tautosyllabic consonant clusters.

Likewise, the vowel inventory of Maasai (Tucker and Mpaayei 1955; Archangeli and Pulleyblank 1994; Baković 2000) is in all relevant respects just like the vowel inventory of Akan described above, but the unpaired low vowel /a/ only blocks leftward [±ATR] harmony; in the rightward direction, /a/ becomes [+ATR] but only by further raising and rounding to become [o]. This raising-and-rounding rule is clearly triggered by the independent absence of a [+ATR] low vowel in the vowel inventory; if the statement of the rule could thereby be reduced just to the result of rightward harmony ("raise and round a vowel that undergoes rightward harmony"), then it would technically underapply in all cases of non-low vowels.

Note that the triggered counterpart of a grandfather effect would simply be any rule the conditions for application of which can be motivated by universal markedness considerations. For example, a rule of syllable-final obstruent devoicing can be and has been argued to be motivated by the relative markedness both of voiced obstruents and of maintaining contrasts in (the rough equivalent of) syllable-final position (Lombardi 1991, 1999; Steriade 1999); this rule might thereby be reduced to the bare minimum "change (obstruent) voicing" – effectively, a rule-based imperative corresponding to a faithfulness constraint in Optimality Theory – and thus underapply when an obstruent is voiceless or not syllable-final.

Aside from the issue of underapplication, do-something-except-when blocking and do-something-only-when triggering are generally anomalous phenomena within rule-based serialism. The logic of these phenomena entails the consideration

of parallel hypothetical derivations at every potential blocking or triggering turn. In order to block a rule from applying to a representation, a hypothetical application of the rule to that representation must be contemplated and found to be in violation of the blocking condition; the result is thereby discarded, and the derivation proceeds without application of the rule. In order to trigger the application of a rule to a representation, a hypothetical *non*-application of that rule must be contemplated and found to be in violation of the triggering condition; this result is thereby discarded and the derivation proceeds with application of the rule.

The necessity of these parallel derivations is rarely if ever acknowledged even in work promoting models that more explicitly acknowledge blocking and triggering (e.g. Paradis 1987; Calabrese 2005; see Odden, this volume, for discussion of some of these kinds of models). Parallel derivations are of course very much like the multiple output candidates of Optimality Theory, which was designed with blocking and triggering phenomena firmly in mind (see especially Chapters 3 and 4 of Prince and Smolensky 1993) and in which the analytical counterpart of any rule necessarily involves forced violations of some constraints; these violations roughly register the various forms of "underapplication" discussed here.

3.2.4 Summary Rules with blocking conditions underapply when they are blocked and rules with triggering conditions underapply when they are not triggered. Since satisfactory accounts of these phenomena require principles or mechanisms beyond rule ordering (the Elsewhere Condition for disjunctive blocking, one of the several proposed accounts of non-derived environment blocking, parallel hypothetical derivations for do-something-except-when blocking and do-something-only-when triggering), rule ordering is clearly insufficient to handle all examples of opacity. (This is of course true regardless of whether the additional principles or mechanisms that these phenomena require are reducible to each other or to something more general.)

3.3 Other Examples of Underapplication

I briefly consider here three additional examples of underapplication as defined in (8): the restriction of a rule to particular lexical classes or levels (Section 3.3.1), rule exceptions (Section 3.3.2), and rule optionality (Section 3.3.3). The identification of at least some of these types of rules as examples of underapplication is not entirely novel: rules that apply only to particular lexical classes and rules that have lexical exceptions fall into the class of "non-automatic" rules, defined by Kiparsky (1976) as those rules for which there are representations *in the immediate output of the rule* – that is, not necessarily on the surface – to which the rule could (still) apply non-vacuously.[10] The classification of at least some of these phenomena as instances of opacity may nevertheless seem counterintuitive, but as I discuss in each subsection, appropriate amendments to the definition of underapplication appear to be nothing but ad hoc. More importantly, recall that the only hypothesis tying opacity together is Kiparsky's claim that instances of it are relatively hard to learn; at a minimum, then, the relative learnability of all of these

phenomena needs to be empirically determined before we begin to write some phenomenon or other in or out of the definition of opacity.

3.3.1 Restriction to Classes/Levels If a given rule applies to some but not all lexical classes or in some but not all levels, then that rule by definition underapplies with respect to the complement set of classes or levels and is thus opaque. For example, the fact that Velar Softening in English (putatively responsible for e.g. *opaque* [k] ~ *opacity* [s]) applies only to the Latinate vocabulary class means that the rule underapplies elsewhere; likewise, the fact that the rule responsible for antepenultimate main stress in English applies at Level 1 (*oríginal ~ originálity*) means that the rule underapplies at later levels (*óbvious ~ óbviousness; *obvíousness*).[11]

If this conclusion seems counterintuitive in the case of lexical classes, one could try to dismiss it by further specifying the denotation of "surface representations" in the definition of underapplication in (8) as the set of representations defined by the particular class to which the relevant rule is restricted to apply. But unless and until we can establish a relevant difference in the relative learnability of class-restricted rules and rules that underapply for other reasons (e.g. because they are counterfed), this move would be completely ad hoc.

This kind of move would not even be desirable in the case of levels because level ordering is generally an accepted mechanism for describing opaque interactions between phonological rules. For example, recall Kiparsky's (1982a) analysis of Shortening and Lengthening in English mentioned in Note 7: underapplication of Shortening is arguably due not to extrinsic within-level ordering nor to disjunctive blocking (by the Elsewhere Condition or otherwise), but rather to the independently-motivated assignments of Shortening to a cyclic level and of Lengthening to a postcyclic level. Some researchers have even claimed that *all* counterfeeding and counterbleeding interactions are due to the (independently-motivated) assignment of different rules to different levels that are serially ordered with respect to each other but within which there is no serial ordering, most notably Kiparsky (to appear) and Bermúdez-Otero (to appear); cf. McCarthy (2007b: 38ff.).

3.3.2 Exceptions If a given rule has (lexical) exceptions, then that rule by definition underapplies with respect to those exceptional forms and is thus opaque. For example, the (independently optional) rule of postnasal /t/ deletion in English (/t/ → ∅ between /n/ and an unstressed vowel; see Hayes 2009: 191–192) exceptionally underapplies in the case of *intonation* for many speakers of English: [ˌĩntə̃'neɪʃə̃n] ~ *[ˌĩnə̃'neɪʃə̃n] (cf. *intellectual* [ˌĩntə'lekʧwəl] ~ [ˌĩnə'lekʧwəl]).

The conclusion that exceptions contribute to opacity is perhaps not so counterintuitive, but it does depend on exactly how rule exceptions are encoded in the grammar and whether the definition of underapplication opacity in (8) is sensitive to that encoding. Much as in the case of lexical class restrictions, any move to redefine underapplication to accomodate exceptions would be ad hoc unless and until a relevant difference in the relative learnability of rule exceptions and rules that underapply for other reasons is established.

3.3.3 Optionality If a given rule is optional, then by definition that rule sometimes underapplies and is thus opaque. For example, consider the optional rule of *t*/*d*-deletion in many varieties of English (see e.g. Coetzee (2004) and references therein): a form like *west* is sometimes realized as [wɛs] and other times as [wɛst]; in the latter case, *t*/*d*-deletion underapplies.

If this conclusion seems counterintuitive, one could again try to dismiss it by redefining underapplication opacity. For example, specifying the "phonological rule ℙ of the form . . ." as "obligatory" would successfully, albeit stipulatively, render optional rules transparent.[12] However, this would also incorrectly exclude cases in which optional rules are uncontroversially opaque not due to their optionality but due to their interaction with other rules; see Kawahara (2002), Anttila (2006), Ettlinger (2007), and Anttila *et al*. (2008) for examples.

Another possibility is to adopt the grammar competition approach to optionality of e.g. Kroch (1989).[13] If each member of a set of possible surface realizations of a given form results from a different grammar, then optionality can be brought into the fold of transparency by saying that a rule ℙ underapplies only if there are surface structures meeting ℙ's input structural description *that are generated by a grammar that includes ℙ*. This is of course a very reasonable amendment to (or clarification of) the definition of underapplication; deliberately excluding it appears to lead to the seemingly absurd but logical conclusion that, in the case of a speaker of two languages L1 and L2, a rule ℙ that is unique to L1 is opaque simply because there are surface structures meeting ℙ's input structural description in L2!

There are two comments that I could make about the seeming absurdity of this logical conclusion. The first comment is that we *do* know that the grammar of one's native language can interfere with the learning of an additional language, and that at least one form of interference involves rules in the native language that do not apply in the additional language (Broselow 1983); moreover, recent research suggests that the process of acquiring multiple native languages may also involve this type of interference (Fabiano-Smith and Barlow 2010).[14] If opacity boils down to relative learnability, as Kiparsky originally suggested, then there appears to be no reason *not* to consider these types of interference between languages as types of opacity. The second comment is that, even granting the grammar competition approach to optionality, there is more than likely a continuum of conscious distinguishability between competing grammars within the same language (= less consciously distinguishable) on one end and non-competing grammars of separate languages (= more consciously distinguishable) on the other – with many points in between, of course. The relative conscious distinguishability of the grammars of separate languages vs. competing grammars within the same language could curtail the impact of opacity in the former case compared to the latter.

These comments stand apart from the by-now-familiar fact that we do not know what differences may or may not exist between the relative learnability of optionality and other forms of underapplication – and multiple language learning, for that matter. As implied throughout the preceding subsections, necessary

empirical work needs to be undertaken before we jump to any conclusions about what should count as opaque and what should not.

3.4 Surface-true Counterfeeding

Another useful term introduced into the discussion of opacity by McCarthy (1999: 332) is SURFACE TRUTH: the generalization expressed by a phonological rule is *not surface true* if there are surface counterexamples to that generalization. The definition of underapplication opacity in (8) technically evaluates the surface truth of a rule, not whether the rule "applies" in all relevant derivations; however, the two notions are sufficiently co-extensive, at least in the simplest case of a pairwise interaction, that "rule ℙ underapplies" and "rule ℙ is not surface true" can be used interchangeably. Here I discuss an example in which a counterfed rule "underapplies" in the narrower sense that it does not apply in a relevant derivation, but in which the generalization expressed by that rule is nevertheless surface true.

In Educated Singapore English (Mohanan 1992; Anttila *et al.* 2008) there are several rules affecting word-final consonant clusters, three of which are discussed here. Epenthesis inserts a schwa between near-identical word-final consonants, much as in standard English (/reɪz+z/ → [reɪz+əz] 'raises'; cf. /bæg+z/ → [bæg+z] 'bags'). Deletion deletes a word-final plosive if it is preceded by an obstruent (/test/ → [tes] 'test'; cf. /test+iŋ/ → [test+iŋ] 'testing'). Finally, Degemination, fed by Deletion, deletes one of two word-final near-identical consonants (/list+z/ $\xrightarrow{\text{Del.}}$ |l$^{\text{D}}$is+z| $\xrightarrow{\text{Deg.}}$ [lis] 'lists').[15]

As Anttila *et al.* (2008: 185) explain, Deletion counterfeeds Epenthesis in the last of these derivations: application of Deletion results in an intermediate representation, |lisz|, to which Epenthesis is applicable, but Epenthesis does not apply; Degemination, which is also applicable, applies instead. Thus Epenthesis must apply before Deletion (= counterfeeding) and Deletion must apply before Degemination (= feeding). But despite the fact that this is counterfeeding, it does not strictly involve underapplication opacity. The fed application of Degemination ultimately removes the structural description of Epenthesis whenever Epenthesis is counterfed by Deletion, the end result being that there are in fact no surface representations to which Epenthesis could apply non-vacuously. Because Epenthesis itself is not responsible for this fact, it "underapplies" – but only in a narrower sense than justified by the definition of underapplication opacity in (8) because Epenthesis is surface true.

On the other hand, if the conspiracy behind Epenthesis and Degemination – to wit, the avoidance of surface (near-)geminates – is factored out of the formal statements of these rules in the way advocated by Kisseberth (1970), then *both* Epenthesis *and* Degemination technically underapply as defined in (8). (See Note 9 and surrounding discussion.) This is consistent with the intuition expressed by Anttila *et al.* (2008: 185) when they state that "[t]he system [of rules affecting consonant clusters in Educated Singapore English – *EB*] exhibits remarkably deep opacity," the counterfeeding interaction between Epenthesis and Deletion being one of five interactions claimed to contribute to this remarkable depth. One of

the others is another counterfeeding interaction between Epenthesis and a rule of Metathesis, which amounts to exactly the same thing as the counterfeeding interaction between Epenthesis and Deletion because Metathesis also ultimately feeds Degemination. (The remaining three interactions are all examples of counterbleeding and are discussed in Section 4.4 further below.)

4 Overapplication and Counterbleeding

The definitions of overapplication opacity in (6b) and of the counterbleeding relation in (3d) are repeated (in suitably modified forms) in (19) and (20), respectively.

(19) A phonological rule $\mathbb{P}$ of the form $A \rightarrow B \ / \ C__D$ OVERAPPLIES if there are surface structures with instances of B derived by $\mathbb{P}$ in environments other than $C__D$.

(20) $\mathbb{B}$ COUNTERBLEEDS $\mathbb{A}$ if $\mathbb{B}$ eliminates potential inputs to $\mathbb{A}$ and $\mathbb{A}$ precedes $\mathbb{B}$.

In Section 4.1 I explain how typical examples of counterbleeding lead to overapplication as defined in (19). In Section 4.2 and Section 4.3 I discuss two types of examples of overapplication that involve (something more like) feeding than counterbleeding, and in Section 4.4 I show that counterbleeding does not always lead to overapplication as defined in (19).

4.1 Counterbleeding

The counterbleeding relation (20) covers situations where an earlier-ordered rule $\mathbb{A}$ applies to a representation that is subsequently changed by a later-ordered rule $\mathbb{B}$ such that the application of $\mathbb{A}$ appears to have been unjustified; $\mathbb{A}$ overapplies in such cases. Consider as an example of both counterbleeding and overapplication the following two rules of Polish.[16]

(21) Counterbleeding in Polish (Bethin 1978; Kenstowicz and Kisseberth 1979)

	a. /ʒwob/	b. /sol/	c. /gruz/
Raising: $\begin{bmatrix}+\text{back}\\-\text{low}\end{bmatrix} \rightarrow [+\text{high}] \ / \ __ \begin{bmatrix}+\text{voi}\\-\text{nas}\end{bmatrix}$	u	u	
Devoicing: [–son] → [–voi] / __ #	p		s
	[ʒwup]	[sul]	[grus]

Glosses: (21a) ‘crib’, (21b) ‘salt’, (21c) ‘rubble’

The derivations in (21b–c) illustrate the independent action of each of the rules: Raising applies alone in (21b) and Devoicing applies alone in (21c), with no

interaction in either case. In (21a), Devoicing counterbleeds Raising because the earlier application of Raising is justified in part by the fact that the following obstruent is voiced, and this critical fact about the context is subsequently changed by Devoicing. Raising thus overapplies because there are raised back round vowels that are not followed by voiced non-nasals on the surface.

The Polish case in (21) is an example of COUNTERBLEEDING ON ENVIRONMENT, because Devoicing crucially changes part of the environment that justified the prior application of Raising. There are also examples of COUNTERBLEEDING ON FOCUS, where both rules affect the same segment as in the following two rules of certain dialects of Low German.

(22) Counterbleeding in Low German (Kiparsky 1968c; Kenstowicz and Kisseberth 1971)

	a. /taːg/	b. /taːg+ə/	c. /haʊz/
Spirantization: $\begin{bmatrix} -\text{son} \\ +\text{voi} \end{bmatrix} \rightarrow$ [+cont] / V __	ɣ	ɣ	
Devoicing: [–son] → [–voi] / __ #	x		s
	[taːx]	[taːɣ+ə]	[haʊs]

Glosses: (22a) ‘day’, (22b) ‘days’, (22c) ‘house’

Spirantization applies alone in (22b) and Devoicing applies alone in (22c). In (22a), Devoicing counterbleeds Spirantization because the earlier application of Spirantization is justified in part by the fact that the to-be-devoiced obstruent is voiced. Spirantization thus overapplies because there are spirantized obstruents on the surface that are not voiced.

Unlike counterfeeding, the distinction between “on focus” and “on environment” here is inconsequential; both are equally problematic or equally unproblematic for theoretical frameworks without (some analog of) serial ordering of phonological operations. Both are problematic for “classic” Optimality Theory, for example (McCarthy 1999, 2007b), and both are equally unproblematic for the Universally Determined Rule Application hypothesis of Koutsoudas *et al.* (1974), in which the rules in (21) and (22) would simply apply simultaneously to the same – in this case, the underlying – representation.

4.2 Self-destructive Feeding

Kiparsky (1971: 612) claims that “the unmarked status of feeding order is not subject to any serious doubt,” meaning that both of Kiparsky’s hypotheses discussed in Section 2 classify feeding as an order-to-be-diachronically-attained since it leads to both maximal utilization and transparency. But as it turns out, there exist types of feeding rule orders that involve overapplication opacity. One type is what I call SELF-DESTRUCTIVE FEEDING, in which an earlier rule feeds a later rule

that in turn crucially changes the string such that the earlier rule's application is no longer justified. An example from Turkish is shown in (23).[17]

(23) Self-destructive feeding in Turkish (Kenstowicz and Kisseberth 1979)

	a. /ajag+sɯ/	b. /ʧan+s ɯ/	c. /bebeg+i/
Elision:			
s/j → Ø / C __	Ø	Ø	
Deletion:			
g → Ø / V __ V	Ø		Ø
	[aja + ɯ]	[ʧan+ ɯ]	[bebe +i]

Glosses: (23a) 'his foot', (23b) 'his bell', (23c) 'baby (acc.)'

Elision applies alone in (23b) and Deletion applies alone in (23c). The derivation in (23a) shows the self-destructive feeding interaction between the two: the result of Elision crucially places the stem-final /g/ in the intervocalic position that causes it to undergo Deletion (that is, Elision feeds Deletion) but the /g/ itself was a necessary part of the environment justifying the application of Elision in the first place (that is, Elision overapplies). This case is an example of SELF-DESTRUCTIVE FEEDING ON ENVIRONMENT, because Deletion crucially changes part of the environment that justified the prior application of Elision; see Baković (2007: 247ff.) for extensive discussion of an example of SELF-DESTRUCTIVE FEEDING ON FOCUS, which – somewhat counter-intuitively – does not involve overapplication.[18]

4.3 Cross-derivational Feeding

Another type of overapplication opacity that is not due to counterbleeding is what I call CROSS-DERIVATIONAL FEEDING. The name is meant to highlight the fact that this kind of feeding interaction cannot be handled within a single derivation; two separate derivations must be considered, one in which the feeding rule creates the conditions for the fed rule to apply in the other derivation. Because the opaque nature of cross-derivational feeding is the main thrust of Baković (2007), I attempt to merely summarize that discussion here.

Cross-derivational feeding can be demonstrated with the well-known example of the past tense alternation in English. Reviewing the facts: the past tense suffix /d/ becomes voiceless after stems ending in voiceless obstruents (e.g. /pæk+d/ → [pæk+t] 'packed') and is separated from the stem by an epenthetic vowel if the stem ends in a near-identical consonant /d/ or /t/ (e.g. /pæd+d/ → [pæd+əd] 'padded', /pæt+d/ → [pæt+əd] 'patted').

The standard analysis of this set of facts, illustrated in (24) below (see Baković 2005: 284ff. for discussion and references), has it that Epenthesis applies between word-final near-identical consonants (that is, word-final consonants that differ at most in voicing), thus applying to both /pæt+d/ (24a) and /pæd+d/ (24b). (Near-identity is loosely represented in the statement of Epenthesis with differing subscripts: $C_i \approx C_j$.) In the case of /pæk+d/ (24c), Assimilation applies to devoice

the past tense suffix consonant. Given that Assimilation could in principle also have applied to /pæt+d/ (24a) if the order between Epenthesis and Assimilation were reversed (as in Educated Singapore English; recall Note 15 but see also Note 20 below), Epenthesis bleeds Assimilation in this derivation.

(24) English past tense alternation (standard bleeding analysis)

	a. /pæt+ d/	b. /pæd+ d/	c. /pæk+d/
Epenthesis:			
$\emptyset \rightarrow$ ə/ C_i __ C_j#		ə	ə
Assimilation:			
[–son] → [αvoi] / [αvoi] __ #			t
	[pæt+əd]	[pæd+əd]	[pæk+t]

Glosses: (24a) 'patted', (24b) 'padded', (24c) 'packed'

This bleeding interaction correctly describes the fact that Epenthesis rather than Assimilation applies in (24a), but at a cost: Epenthesis must arbitrarily ignore the difference in voicing between the stem-final /t/ and the suffix /d/ – precisely the difference that would be neutralized by Assimilation were it to apply. This redundancy can be eliminated by making *strict identity* a requirement on Epenthesis (again, as in Educated Singapore English) and relying on Assimilation to provide the necessary context in (24a).

But of course Assimilation does not *actually* apply in (24a); it only *potentially* applies, but this potential appears to be sufficient to "feed" the application of Epenthesis instead. A reasonable way to model this type of interaction is with two parallel derivations, one in which Assimilation applies and another in which Epenthesis applies, as shown in (25).

(25) English past tense alternation (cross-derivational feeding analysis)

Assimilation: [–son] → [αvoi] / [αvoi] __ # Epenthesis: $\emptyset \rightarrow$ ə/ C_i __ C_i#

a.	b.	c.
/pæt+d/	/pæd+d/	/pæk+d/
A, E	A, E	A
|pæt+t|—$\mathcal{F}$ → [pæt+əd]	|pæd+d|—$\mathcal{F}$ → [pæd+əd]	[pæk+t]

Assimilation is stated just as in (24) above, but Epenthesis is now stated to apply only between strictly identical word-final consonants ($C_i = C_i$). The idea here is that Epenthesis applies if and only if its structural description is met by the potential output of Assimilation; this is the case in (25a,b) – though vacuously so in (25b) – and so Epenthesis applies to those two examples. It is not the case in (25c), however, and so Assimilation applies in that example. Because the application of Epenthesis in (25a) is motivated only by the potential but not actual non-vacuous application of Assimilation, Epenthesis overapplies in this derivation in accordance with the definition of overapplication opacity in (19).

As discussed in Baković (2005, 2007), the kind of interaction illustrated in (25a) is impossible to describe with the single derivation characteristic of rule-based

serialism because the potential derivation with Assimilation applying is necessary to trigger Epenthesis in the actual derivation, leading to the correct surface representation. This is in fact what makes the bleeding analysis in (24a) a necessary evil, with the arbitrary and redundant stipulation that voicing is the one feature that can be ignored in the determination of near-identity for the purposes of Epenthesis application. Cross-derivational feeding is thus yet another example of an opaque interaction that cannot be accounted for by rule ordering alone.

4.4 Mutual Bleeding

The term MUTUAL BLEEDING, following Kiparsky (1971: 600), refers to situations where a rule 𝔸 bleeds a later-ordered rule 𝔹 and where 𝔹 would also bleed 𝔸 if 𝔹 were ordered before 𝔸.[19] Whether this means that 𝔹 counterbleeds 𝔸 depends on the interpretation of the clause "𝔹 eliminates potential inputs to 𝔸" in the definition of counterbleeding in (20). The fact that 𝔸 precedes and bleeds 𝔹 in a mutual bleeding situation means that 𝔹 does not get to apply in derivations where 𝔸 applies, so there's no opportunity for 𝔹 to *actually* eliminate potential inputs to 𝔸. But if the definition is interpreted more broadly to mean that 𝔹 *in principle* eliminates potential inputs to 𝔸, then mutual bleeding counts as what we might call BLED COUNTERBLEEDING (recall 'fed counterfeeding' from Section 3.1).

Indeed, counterbleeding is often defined to more obviously encompass mutual bleeding; consider for example the following representative textbook definition.

(26) Counterbleeding (adapted from Hayes 2009: 185)

Rule 𝔹 counterbleeds rule 𝔸 when
- 𝔹 is ordered after 𝔸, and
- 𝔹 would have removed configurations to which 𝔸 applies, had 𝔹 applied first.

The "would have removed" part is the key to the inclusion of mutual bleeding, and in fact Hayes uses the following example of mutual bleeding from Lardil to illustrate counterbleeding.

(27) Mutual bleeding in Lardil (Hale 1973; Hayes 2009)

		a. /papi+ uɻ/	b. /tʲæmpæ+uɻ/
Epenthesis:	Ø → w / i __ u	w	
Elision:	V → Ø / V __		Ø
		[papi+wuɻ]	[tʲæmpæ+ ɻ]

Glosses: (27a) 'father's mother (acc. fut.)', (27b) 'mother's father (acc. fut.)'

The derivation in (27b) illustrates the independent action of Elision: the first vowel in hiatus is not an /i/, and so the second vowel is elided. In (27a), Epenthesis bleeds Elision because insertion of the glide separates the vowels in hiatus. Elision thus also counterbleeds Epenthesis here, according to the definition in (26): elision

of the suffix vowel would have removed the necessary /u/ from the context of Epenthesis. A third example illustrating the "independent" action of Epenthesis is impossible to provide, given that Epenthesis applies to a proper subset of cases to which Elision is applicable. Despite its relevance in this case, note that the Elsewhere Condition (Section 3.2.1) is not needed to block Epenthesis when Elision applies because the bleeding relation between the two rules does the trick, but Koutsoudas *et al.* (1974: 8ff.) do propose that such pairs of rules are intrinsically ordered with respect to each other by the related Proper Inclusion Precedence Principle (Sanders 1974).

There is also a mutual bleeding interaction between Epenthesis and Degemination in Educated Singapore English, when the intervening Deletion rule is not involved (recall the interaction among these rules discussed in Section 3.4): /reɪz+z/ → [reɪzəz], *[reɪz]. Epenthesis clearly bleeds Degemination here by separating the members of the would-be geminate. Anttila *et al.* (2008: 185), apparently assuming the definition of counterbleeding in (26), state that Degemination also counterbleeds Epenthesis: had it applied, Degemination would have removed one of the two halves of the geminate from the context of Epenthesis.

Note that these are examples of MUTUAL BLEEDING ON ENVIRONMENT: each rule crucially disrupts the environment required for the application of the other. There are also cases of MUTUAL BLEEDING ON FOCUS, for example the following case from two different sets of dialects of German (Vennemann 1970, Kiparsky 1971: 600). In one set of dialects, the Devoicing rule already discussed in (22) bleeds a Deletion rule that deletes /g/ after nasals: /laŋg/ → [laŋk], *[laŋ] 'long (masc.)'; in the other set of dialects, the order is reversed so that Deletion bleeds Devoicing: /laŋg/ → [laŋ], *[laŋk] (cf. /laŋg+ə/ → [laŋ+ə] 'long (fem.)' in both sets of dialects, given the inapplicability of Devoicing in this case).

Mutual bleeding interactions like these obviously do not involve overapplication. Because Epenthesis bleeds Elision in Lardil, Elision does not get a chance to change the environment that justified the prior application of Epenthesis; in other words, Epenthesis in Lardil does not overapply. Likewise, because Epenthesis bleeds Degemination in Educated Singapore English, Degemination does not get a chance to change the environment that justified the prior application of Epenthesis; thus there are in fact no surface representations to which Epenthesis in Educated Singapore English has overapplied.[20] Finally, because Devoicing bleeds Deletion in some dialects of German and Deletion bleeds Devoicing in others, the bled rule does not change the environment that justified the application of the bleeding rule and so the bleeding rule does not overapply. To the extent that counterbleeding encompasses mutual bleeding, then, not all cases of counterbleeding involve overapplication.

Note that the rules involved in some examples of mutual bleeding can be implicated in a conspiracy. Epenthesis and Elision in Lardil are both hiatus-avoidance strategies, and as already noted in Section 3.4, Epenthesis and Degemination in Educated Singapore English are both (near-)geminate-avoidance strategies.[21] As discussed in Section 3.2.3, factoring out what is being avoided from the structural descriptions of the rules involved in a conspiracy inevitably

results in underapplication; to the extent that mutual bleeding involves counterbleeding, then, we can conclude that some cases of counterbleeding lead to underapplication opacity.

5 Concluding Remarks

The phonology of a language is a complex system, generating a set of surface forms the ultimate token realizations of which serve as the input that language learners are exposed to and presumably use to acquire the system. To the extent that this system is composed of individual phonological rules, it is not unreasonable to assume that the easier it is to isolate the operation of those individual rules from the input, the easier it is to acquire those rules and hence the system. But phonological rules do not generally operate in isolation, nor do they tend to interact in simple pairwise ways. Although phonologists often find it useful, for expository or pedagogical purposes, to (attempt to) isolate the operation of a single phonological rule or the interaction between two rules, it is always important to be mindful of the overall system. Could the actions of other rules affect any conclusions drawn from an individual rule or interaction between rules? Could attention to other parts of the system be necessary to understand the workings of an individual rule or interaction? In the absence of solid answers to these types of questions, we have little basis beyond Kiparsky's suggestive – but by no means conclusive – diachronic evidence that it is hard to learn opaque rules; after all, such questions presumably apply not only to a phonologist's analysis of the phonology of a language but also to a learner's acquisition of one.

The resurgence of research on phonological opacity over the past 15 years or so has unfortunately not paid attention to such questions; opacity has instead been wielded as a weapon in the larger debate between proponents of rule-based serialism and proponents of alternative theoretical frameworks, Optimality Theory in particular. The debate has been sharply polarized in most respects, but there is one mistaken "fact" on which nearly all researchers on both sides (e.g. Vaux 2008, McCarthy 2007b) mysteriously appear to have decided to agree: that rule-based serialism, via its central principle of rule ordering in (2), offers a unique and unified account of opacity as originally defined by Kiparsky in (1). I have demonstrated in this chapter that this is simply not the case, unless we decide to depart from Kiparsky's agreed-upon definition of opacity and instead stipulatively (and perversely) define it as just those opaque interactions that can be described with rule ordering. Further discussions of the implications of opacity for theoretical framework comparison should either acknowledge this or provide a different, principled definition of opacity on which to base such discussions (see e.g. Bermúdez-Otero 1999, Idsardi 2000, Ettlinger 2008, and Tesar 2008, forthcoming).

This result of this demonstration is neither surprising nor a matter of concern. Kiparsky's learnability claim is really all that warrants the investigation of "opacity" as a singular notion, and there is no a priori reason to assume that the relative learnability of a phonological generalization should be reflected in the

formal mechanisms used to account for its interaction with another phonological generalization that is responsible for that relative learnability, and there is even less reason to assume that any two generalizations with similar degrees of learnability should be accounted for with the same formal mechanisms. Even if there were reasons to make such assumptions, there is precious little (if any) research quantifying the relative learnability of different phonological generalizations as a function of their interactions with other phonological generalizations. In the absence of such crucial empirical work, any formal assumptions we make about opacity are bound to be tentative at best.

ACKNOWLEDGMENTS

For stimulating discussions of and helpful comments on the contents of an earlier version of this chapter, I thank the Crosslinguistic Investigations in Syntax-Phonology group at University of California–Santa Cruz (esp. Ryan Bennett and Vera Gribanova), the San Diego Phonology Interest Group at University of California–San Diego (esp. Sharon Rose), the members of Gene Buckley's seminar on opacity at Penn (esp. Josef Fruehwald), the editors of the *Blackwell Companion to Phonology* (esp. Marc van Oostendorp and Beth Hume), the editors of the *Handbook of Phonological Theory* (esp. Jason Riggle), and two anonymous reviewers. Remaining errors are mine to keep.

NOTES

1 Kiparsky (1968c) was one of the first to explicitly distinguish between these relations (see also Chafe 1968; Wang 1969; Koutsoudas *et al*. 1974), and was certainly the first to use the feeding/bleeding terminology; Newton (1971) appears to have introduced the "counter-" prefix. (Kenstowicz and Kisseberth 1971 used a "non-" prefix in the same sense; in later work, Kenstowicz and Kisseberth 1977, 1979 used "counter-.")

2 Albright and Hayes (this volume) discuss an actual example of counterfeeding and counterbleeding in different derivations, arising from the ordering of height-dependent rounding harmony before high vowel lowering in Yokuts (Newman 1944; Kuroda 1967; Kisseberth 1969; Kenstowicz and Kisseberth 1979).

3 For example, one can ask: do the definitions in (3) allow for the possibility that $\mathbb{A}$ feeds $\mathbb{B}$ because $\mathbb{A}$ bleeds some intervening rule $\mathbb{C}$ that would otherwise bleed $\mathbb{B}$? (And: is the answer intuitively correct?)

4 See also recent work by the research team associated with the Learnability Project at Indiana University (e.g. Barlow 2007; Part II of Dinnsen and Gierut 2007; Dinnsen and Farris-Trimble 2008), which documents cases of opacity that appear to arise spontaneously during the course of language acquisition. Vaux (this volume) also notes examples of spontaneous opacity arising in language games.

5 The usefulness of these terms in describing the often special phonology of reduplication was highlighted by McCarthy and Prince (1995, 1999) and was first adapted to other phenomena by Benua (1997); see Section 3.3.1.

6 The "$\sigma\ \sigma$" in the environment of Apocope is meant to denote the fact that the rule is blocked from creating monosyllabic words (Wilkinson 1988; Prince and Smolensky 1993), and the ad hoc feature [–apical] denotes the disjoint set of [–coronal] and [+distributed] consonants that are targeted by Deletion.

7 In Kiparsky (1982a: 154ff.), Shortening is independently classified as a cyclic rule (because it is blocked in non-derived environments; see Section 3.2.2) while Lengthening is independently classified as a postcyclic rule. Lengthening is thus *intrinsically* ordered after Shortening in this analysis; see Section 3.3.1 for more on this point.

8 See McCarthy (2002: 63) for a comprehensive bibliography of 1970s-era work on conspiracies.

9 Kiparsky (1976: 80ff.) comes to the opposite conclusion about conspiracies, stating that "the fact that languages tend to have conspiracies follows from the more general fact that languages tend to have transparent rules." This conclusion comes on the heels of an argument against Kisseberth's proposal that the rules participating in a conspiracy should have the function of the conspiracy factored out of their statements, Kisseberth's claim being that this formally simplifies the grammar. Kiparsky argues that Kisseberth's invocation of Chomsky and Halle's (1968) formal evaluation metric is unsuccessful, but given the general lack of success of the evaluation metric – at least in the crude, feature-counting form that is relevant to the discussion – this argument does not necessarily undermine Kisseberth's underlying proposal.

10 Optional rules are also non-automatic unless analyzed in terms of grammar competition; see Section 3.3.3.

11 Indeed, Benua (1997) adapts the terms "underapplication" and "overapplication" from Wilbur (1973) to describe just these sorts of differences in rule applicability in different levels; recall Note 5.

12 Note that the Revised Alternation Condition of Kiparsky (1976), noted briefly in Section 3.2.2, likewise stipulates that only *obligatory* neutralization rules are blocked from applying in non-derived environments.

13 I am indebted to Josef Fruehwald for raising the issues discussed in the remainder of this subsection.

14 Thanks to Cynthia Kilpatrick and Bożena Pająk for the representative references cited here.

15 In Mohanan's analysis, Degemination only applies to clusters of *strictly identical* consonants and must thus also be fed by a voicing assimilation rule not discussed here (/list+z/ $\xrightarrow{\text{Del.}}$ |lis+z| $\xrightarrow{\text{Assim.}}$ |lis+s| $\xrightarrow{\text{Deg.}}$ [lis]). The simplification in the text does not affect the point at issue; see Section 4.3 for more relevant discussion.

16 See Buckley (2001) and Sanders (2003) for an alternative view of the Raising alternation.

17 See also Inkelas, this volume, where Paster's (2006: 99) input subcategorization analysis of this example is summarized. (Thanks to Jorge Hankamer for instructing me on the finer points of the Deletion rule.)

18 Self-destructive feeding was first identified as an opaque feeding order in Baković (2007); the example of NON-GRATUITOUS FEEDING discussed in that article is left out here in the interests of space.

19 Thanks to Marc van Oostendorp for very helpful comments on the content of this section.

20 The mutual bleeding interaction in Educated Singapore English between Deletion and Metathesis also does not involve overapplication for the same reasons. In the end, only one of the five interactions contributing to the "remarkably deep opacity" of this system – counterbleeding between Epenthesis and Voicing Assimilation, mentioned

in Note 15 – is in fact an opaque one according to Kiparsky's definition in (6), and is the one interaction that Anttila *et al.* (2008: 194ff.) ultimately deny the factual basis of. Kenstowicz and Kisseberth (1971) put forth the idea that counterbleeding interactions between epenthesis and assimilation rules might be universally non-existent; see also Baković (2007: 245ff.) and Baković and Pająk (2008).

21 The German case is at best a less-than-clear example of a conspiracy. Both Deletion and Devoicing can do their part in ridding the surface of [ŋg] sequences, but both only actually do so in the first set of dialects described in the text; in the second set of dialects, Devoicing never gets a chance to apply to the relevant sequences. Furthermore, Devoicing more generally devoices all syllable-final obstruents, not just /g/.

3 The Interaction Between Morphology and Phonology

SHARON INKELAS

1 Introduction

The morphology of a language concerns the generalizations about form and meaning that relate words to one another within that language. The phonology of a language concerns the generalizations about the sound patterns in that language. Morphology and phonology intersect insofar as the statement of morphological generalizations includes information about sound patterns, or insofar as the statement of phonological generalizations includes information about morphology.

2 When Morphology Affects Phonology: The Phonological Interpretation of Morphologically Complex Words

The earliest influential generative approaches to the intimate interaction between phonology and morphology (Chomsky and Halle 1968; Kiparsky 1982b; Mohanan 1986) focused on the phonological interpretation of morphologically complex words, and this is where we will begin our survey as well, although we will not restrict ourselves to the phenomena covered by any particular theory in the process.

2.1 Morphologically Conditioned Phonology

Phonological requirements in a language can alter the shape that individual morphemes take in different contexts, producing allomorphy. Sometimes these

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

alternation patterns are quite general in the language. In Turkish, for example, a very general rule of progressive vowel harmony determines the value of [back] for the vowels of most suffixes, which surface with front vowels following roots whose final vowel is front (e.g. *gyl-ler* 'rose-PL', *anne-ler* 'mother-PL') but with back vowels following stems whose final vowel is back (e.g. *ok-lar* 'arrow-PL', *elma-lar* 'apple-PL') (see, for example, Lewis 1967). Morphologically conditioned phonology arises when phonological alternations are *not* fully general in the language but are instead specific to particular morphological constructions, such as compounding, truncation, affixation, or reduplication (for overviews at a fairly theory-neutral level, see, for example, Dressler 1985; Spencer 1998).

In Belhare, for example, intervocalic voicing occurs at stem-suffix boundaries (*lap* > *lab-u!* 'catch it!') but not at prefix-stem boundaries (*ka-pira!* 'give it to me!'), or in underived words (*pipisi* '(drinking) straw') (Bickel and Nichols 2007). In Turkish, the diminutive suffix *-CIk* triggers the deletion of stem-final *k* (Lewis 1967: 57): *bebek, bebe-cik* 'baby/baby-DIM', *köpek, köpe-cik* 'dog/dog-DIM'). No other consonant-initial suffix triggers this deletion (*bebek-çi* 'child care provider', *bebek-lik* 'infancy', *bebek-ten* 'baby-ABL', *bebek-ken* 'while a baby', etc.). In Dakota, a coronal → velar dissimilation rule targets coronal consonant clusters that straddle the juncture between the two copies in reduplication (/žat/ → *žag-žát-a* 'curved'), but not clusters arising in other morphological contexts, for example, compounding (*sdod* + *čhí-ya* 'know.I + you-cause = I know you'; Shaw 1985: 184).

Morphologically conditioned phonology can be segmental, as in the examples just cited, or can involve prosodic properties such as tone, stress, or length. Very familiar examples include Indo-European accentuation (Kiparsky 1973b) and Japanese (McCawley 1968b; Poser 1984; Pierrehumbert and Beckman 1988; Alderete 1999, 2001). In Japanese, morphological constructions, which include prefixation, suffixation, zero-derivation, and compounding, come in two essential varieties: those which preserve lexical stem accent and those which erase it. Poser (1984) terms the two types "recessive" and "dominant," respectively, building on terminology introduced in Kiparsky 1973b (see also Kiparsky and Halle 1977; Halle and Mohanan 1985). Japanese pitch-accent is subject to strict distributional regularities: each word has at most one accent, and in cases of conflict between two lexically accented morphemes in the same word, the general principle is that the leftmost accent wins (Poser 1984). Recessive suffixes, as shown in (1), behave according to the Leftmost Wins principle. An unaccented suffix, for example, past tense *-ta*, leaves stem accent unaffected (1a), while an accented recessive suffix, for example, conditional *-tára*, surfaces with its accent only if the stem is not already lexically accented (1b). Otherwise, Leftmost Wins results in the elimination of suffix accent (1c). Page numbers are from Poser (1984):

(1)	a.	/yob-ta/	→	yoNda	'called'	(49)
		/yóm-ta/	→	yóNda	'read'	(49)
	b.	/yob-tára/	→	yoNdára	'if he calls'	(48)
	c.	/yóm-tára/	→	yóNdara	'if he reads'	(48)

Other recessive suffixes are pre-accenting, depositing accent on the final syllable of unaccented stems but having no effect on lexically accented stems (2a). Still others are accent-shifting. Poser terms these "dependent"; they shift stem accent, if any, to the stem-final syllable, but do not have any effect on lexically unaccented stems (2b):

(2)	a.	matumoto-si	→	matumotó-si	'Mr Matsumoto'	(54)
		áNdoo-si	→	áNdoo-si	'Mr Ando'	(54)
		nisímura-si	→	nisímura-si	'Mr Nishimura'	(54)
	b.	koná-ya	→	konáya	'flour seller'	(55)
		kúzu-ya	→	kuzúya	'junk man'	(55)
		kabu-ya	→	kabuya	'stockbroker'	(55)

In contrast to recessive affixes, dominant affixes trigger deletion of stem accent. Accented dominant suffixes, like adjective-forming *-ppó*, erase stem accent and surface themselves as accented (3a). Unaccented dominant affixes produce completely unaccented outputs, like demonymic *-kko* (3b). Still other dominant suffixes place accent on the initial or final stem syllable, as illustrated by (most forms with) the "true" prefix *ma(C)-* (3c) and family naming *-ke* suffix (3d), or even on the stem-penultimate syllable, as with the girls' name-forming *-ko* (3e):

(3)	a.	abura	→	abura-ppó-i	'oil, fat/oily'	(49)
		yásu	→	yasu-ppó-i	'cheap/cheap, tawdry'	(49)
		adá	→	ada-ppó-i	'charming/coquettish'	(49)
	b.	kóobe	→	koobe-kko	'an indigené of Kobe'	(72)
		nágoya	→	nagoya-kko	'an indigené of Nagoya'	(72)
		nyuuyóoku	→	nyuuyooku-kko	'an indigené of New York'	(72)
	c.	futatu	→	map-pútatu	'two/exactly half'	(57)
		sáityuu	→	mas-sáityuu	'amidst/in the very midst of'	(57)
		syoozíki	→	mas-syóoziki	'honesty/downright honest'	(57)
	d.	nisímura	→	nisimurá-ke	'the Nishimura family'	(55)
		ono	→	onó-ke	'the Ono family'	(55)
		hára	→	hará-ke	'the Hara family'	(55)
	e.	haná	→	hána-ko	'flower/name'	(58)
		kaede	→	kaéde-ko	'maple/name'	(59)
		mídori	→	midóri-ko ~ midorí-ko	'green/name'	(59)

Thus for each affix, or more generally for each morphological construction, since zero-derivation and compounding are subject to similar accentual parameters, it is necessary to know which of several possible accent placement patterns the construction triggers (none, stem-initial, stem-final, stem-penultimate) and whether those patterns preserve or delete lexical stem accent (dominant vs. recessive).

A more unusual case occurs in the Mayan language Mam (England 1983; Willard 2004), in which vowel length is contrastive both in roots and in suffixes

and at most one long vowel is permitted per word. Suffixes divide into two types: those that trigger shortening of stem vowels, and those that do not. Willard terms these "dominant" and "recessive" suffixes, respectively, following the terminology used in the accentual literature. Vowel length of the suffix itself is not a predictor of vowel shortening, as shown in the table; neither is stress nor morphological function. Whether or not suffixation causes stem vowel shortening is an idiosyncratic property of each suffixation construction. Dominant suffixes are shown in (4a); recessive suffixes are shown in (4b):

(4)	a.	mool-	'burn'	mol-oloon	'easily wilted' (FACILITATIVE)
		juus-	'burn'	jus-b'een	'burned place' (RESULTANT LOCATIVE)
		jaaw-	'go up'	jaw-nax	'up' (DIRECTIONAL)
		yuup-	'put out fire'	yup-na	'put out' (PARTICIPIAL)
	b.	iil-	'sin'	iil-a	'scold' (INTRANSITIVE VERBALIZER)
		ooq'-	'cry'	ooq'-b'il	'something which causes crying' (INSTRUMENTAL)

While cases of stress and tone replacement are more common than vowel length manipulation, on the basis of current knowledge it seems reasonable to assume that any kind of phonological pattern, other than the most low-level allophonic alternations, can be restricted to a morphological context, in some language or another. Indeed most phonetically "unnatural" phonological alternations (see e.g. Anderson 1981; Buckley 2000, Hyman 2001a) are morphologically conditioned in just this way, maintaining their niche of productivity in specific morphological contexts (see e.g. Pierrehumbert 2006b).

How is morphologically conditioned phonology to be handled? Current thinking, building on ideas going back to the 1960s, offers two main options: cophonologies, which are co-existing sub-grammars within a single language, each indexed to a particular morphological construction or set of constructions (e.g. Orgun 1996; Anttila 2002a; Itô and Mester 1995; Inkelas and Zoll 2005); or indexed constraints, in which the language has just one phonological grammar, but particular constraints within it are indexed to specific morphemes or morphological constituents (e.g. McCarthy and Prince 1995; Itô and Mester 1999; Smith 1999; Alderete 2001; Pater 2009).

To handle Japanese accentuation, for example, a cophonological approach would subdivide the grammar into a number of closely related variants, and index each morphological construction to one of these variants (cophonologies). The "dominant" morphological constructions would be associated with cophonologies in which input stem accent is eliminated. Poser (1984), using a rule-based precursor to cophonologies, proposed indexing an accent deletion rule to each dominant affix. In an Optimality Theory (Prince and Smolensky 2004; McCarthy 2008) implementation of cophonology theory, the same goal would be accomplished by varying the ranking of the constraints characterizing a particular accentuation pattern either below or above the faithfulness constraint, preserving stem accent. For example, suppose stem-final and stem-initial accent are imposed by the

alignment constraints ALIGN-RIGHT(accent, stem) and ALIGN-LEFT(accent, stem), respectively. The cophonology of a dominant suffix would rank its accent-placing constraints above MAX-accent, ensuring the deletion of stem accents that are in the wrong location. The cophonology of a recessive affix would rank MAX-accent highest, ensuring that the relevant accentuation pattern is imposed only as a default.

All affix cophonologies in Japanese share the constraint ranking, ensuring that there is at most one accent possible in the output. Generally, in cophonological models, the great majority of constraint rankings are shared by all cophonologies in the language; Anttila (2002a) has modeled this sharing using an inheritance hierarchy, in which cophonologies are grouped together by the constraint rankings that unite them. The superordinate node in such a hierarchy, or what Inkelas and Zoll (2005) term the "master ranking," represents the unique genius of the language, a partial ranking of constraints to which every individual cophonology must conform. As Anttila's work makes clear, it is also possible to group smaller subsets of cophonologies under intermediate nodes to capture subregularities, for example, overall differences between nouns and verbs in Japanese (as documented by McCawley 1968b and Poser 1984) or between nouns and adjectives in Finnish (Anttila 2002a).

Constraint indexation is a different, contemporary approach to morphologically conditioned phonology, developed in the early days of Optimality Theory. The approach was originally morpheme-based, indexing constraints to particular (sets of) morphemes. For instance, Itô and Mester (1999) account for the resistance of recently borrowed roots in Japanese to native phonotactic restrictions such as No-P, the ban on [p], by indexing a special, high-ranked faithfulness constraint to exactly the set of relevant roots: $\text{FAITH}_{\text{AssimilatdForeign}}$ >> No-p >> $\text{FAITH}_{\text{Yamato}}$. A native root with underlying illicit /p/ would have to get rid of /p/, but an assimilated foreign root, as in *pato-kā* 'patrol car' (p. 63), would preserve it. Constraint indexation has also been applied to derived stems, nearly merging the difference between cophonologies (indexed to stem-forming constructions) and indexed constraint theory. In his analysis of morphologically conditioned accentuation in Japanese, for example, Alderete (2001) differentiates dominant and recessive affixes by indexing anti-faithfulness constraints to stems derived by the former. The constraint $\neg\text{OO}_{\text{Dom}}$-MAX-ACCENT ("It is not the case that every accent in S1 has a correspondent in S2") specifies that in derived stems created by dominant affixes, an input stem accent is not preserved in output. In Alderete's model, if a dominant affix causes input stem accent to delete, then the default accentuation pattern of the language is imposed in its place; it is also possible for alignment constraints to locate accent at the boundaries between stems and specific affixes.

With regard to the types of substantive differences that can exist between morphophonological patterns in the same language, the two approaches are very similar substantively, at least when implemented in Optimality Theory; each uses the same set of constraints and thus predicts the same range of possible

morphologically conditioned phonological effects. A more probative question is what degree of difference can exist across different morphologically conditioned patterns in the same language. Proponents of indexed constraints have suggested that the bulk of language-internal variation can be described in terms of relative faithfulness. Alderete (2001) has termed this "grammar dependence," claiming that each language has a single set of phonological restrictions (syllable structure, accentuation, segment inventory, etc.); individual morphological constructions differ only in the degree to which they are faithful to input structures which violate these restrictions. Thus, for example, in Japanese, stems created by recessive affixes are faithful to input accent, while stems created by dominant affixes are not faithful, and exhibit the default accentuation pattern of the language. The theoretical arguments on this point are subtle and complex; see, for example, Itô and Mester (1999), Inkelas and Zoll (2007), and Pater (2009). The empirical issue is whether any language ever imposes completely contradictory patterns in different morphological environments. Japanese accentuation is arguably a case of this kind, since even within the set of dominant affixes, at least four contradictory accentuation patterns are observed, as seen in (3).

A related substantive question about morphologically conditioned phonology is the number of variants (cophonologies, indexed constraints) a single language can permit, and the degree of differences among them. This question has been addressed explicitly in work by Anttila (2002a), whose hierarchical cophonological model predicts that every constraint ranking possibility not excluded in the "master ranking" of a language is expected to be instantiated in some cophonology. Neither cophonology theory nor indexed constraint theory addresses the question of how many different cophonologies are possible, or, really, to what degree they could potentially differ. As observed by Itô and Mester (1999), Inkelas and Zoll (2007), and Pater (2009), these issues may ultimately be laid at the feet of the historical origins of cophonological variation, which include language-internal factors like grammaticalization and analogy as well as external factors like lexical borrowing or more extreme language contact, as well as influences of language acquisition.

One way in which the cophonological and indexed constraint approaches clearly differ is in their ability to capture the interaction between different morphologically conditioned patterns in the same language, or "layering effects."

2.2 *Layering Effects*

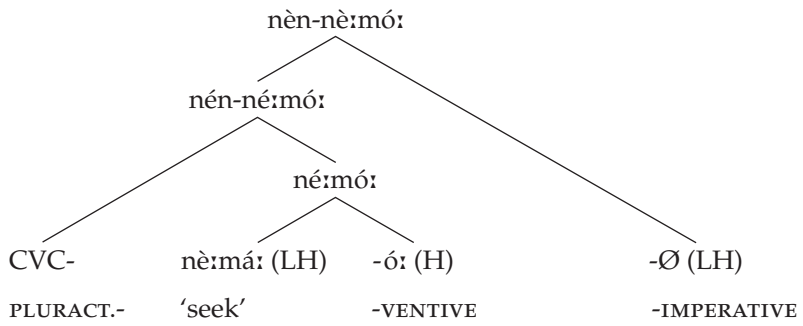
If two morphological constructions are present in the same word, and each is associated with its own phonological pattern, which pattern prevails, or if both do, how do they interact?

The evidence suggests that both patterns prevail, and that they are imposed in the order in which the associated morphological constructions are combined. This is perhaps easiest to illustrate using accentuation patterns that are incompatible,

such that when two morphological constructions affiliated with incompatible patterns co-occur in the same word, one must take precedence over the other. A good case study is Turkish, which resembles Japanese in some of its overall accentuation principles. The default position for stress in Turkish words is final (thus *arabá* 'car', *araba-lár* 'cars', *araba-lar-dán* 'from car'); there is exactly one stress per word, regardless of morphological complexity. A number of morphological constructions assign stress; these always override the default final stress pattern. (On Turkish stress, see e.g. Lewis 1967; Sezer 1981; Kabak and Vogel 2001; Inkelas and Orgun 2003.) In words with more than one stress-assigning morphological construction, order of morphological combination predicts the stress outcome. For example, Turkish has a productive zero-derivation construction forming place names out of words of any part of speech; the construction is marked by a distinctive stress pattern (Sezer 1981) which places stress on the penultimate or antepenultimate syllable, depending on syllable weight: *bak-acák* 'look-FUT' ~ *Bakácak* (place name), *torba-lí* 'bag-ASSOC' ~ *Tórbalı* (place name), and so on. Turkish also has pre-stressing suffixes like past tense predicative -(y)DI (*torbá-ydı* 'it was a bag'), negative -mE (*gel-dí* 'came' vs. *gél-me-di* 'didn't come'), or mitigative -CE (*süt-lü-lér* 'milk-ASSOC-PL = the milky ones', vs. *süt-lű-ce* 'milk-ASSOC-MIT = kind of milky'). As documented in Inkelas (1999), Inkelas and Orgun (1998), Inkelas and Orgun (2003), the stress patterns of the language are all recessive in the sense that they are imposed only if the input stem lacks stress. In words like /torba-lI/, the stress outcome depends on whether an unstressed root, for example, /torba/ 'bag' is first converted to a stressed place name (*Tórba*) and then suffixed (→ *Tórba-lı*), retaining its place name stress, or first suffixed (*torbá-lı*) and then converted to a place name (*Torbálı*), retaining the stress assigned by the suffix instead of displaying the place name stress pattern. Like the Indo-European cases discussed by Kiparsky (1973b), Turkish respects a principle of "Innermost Wins" (Inkelas 1999; Inkelas and Orgun 2003).

Another useful illustration of layering can be found in Hausa, a lexical tone language whose morphological constructions either preserve stem tone (comparable to the "recessive" morphology of Japanese) or replace it with a new tone melody ("dominant") (Newman 1986; 2000; Inkelas 1998). The structure in (5) illustrates a verb root which combines with the dominant ventive suffix *-oː*, then undergoes pluractional reduplication, and is finally converted, via zero-derivation, to an imperative. Both the ventive and the imperative constructions are dominant. The ventive imposes an all-H melody (e.g. *fìtáː* (LH) 'go out' → *fít-óː* (H) 'come out', *gángàráː* (HLH) 'roll down' → *gángár-óː* (H) 'roll down here', and so on (Newman 2000: 663). The imperative imposes a LH melody (e.g. *káːmàː* (HL) → *kàːmáː* (LH) 'catch!', *bíncìkéː* (HLH) → *bìncìkéː* (LH) 'investigate!'; *néːmóː* (H) → *nèːmóː* 'seek!', *nánnéːmóː* (H) → *nànnèːmóː* (LH) 'seek repeatedly!'). In (5), the ventive occurs hierarchically inside the imperative. Predictably in Hausa, the outermost dominant construction is the one whose pattern surfaces; in this case the outermost construction is the imperative, and consequently the whole word surfaces LH. Zero-derivation constructions are represented by null suffixes for purely graphical convenience:

(5)



This kind of pattern is challenging for indexed constraint theory, in which all constraints, morphologically indexed and general, exist in one fixed ranking in the grammar of the language. In Turkish, the constraints that require place names to have the Sezer stress pattern must rank either below or above the constraints requiring stress to immediately precede suffixes like /-lI/. In a word containing both a zero-derived place name and a pre-stressing suffix, the higher-ranked pattern should always prevail, regardless of morphological structure. The problem is that both types of embedding can occur in Turkish, with different meanings and different stress outcomes corresponding to the two possible hierarchical structures (Inkelas and Orgun 1998). A single ranking, as in indexed constraint theory, can capture one but not the other, missing the connection between morphological embedding and constraint ranking. By contrast, in cophonology theories this connection is captured intrinsically (see e.g. Inkelas 1993; Orgun 1996 on "deriving cyclicity"); the hierarchical relationship between two constructions directly determines the input-output relationship between the associated cophonologies.

Some layering theories have bundled layering with additional claims, and have been weakened insofar as the additional claims have not held up. For example, the theory of Lexical Morphology and Phonology (Kiparsky 1982; Mohanan 1986) associated cyclicity (layering) with structure preservation and strict level ordering, to which subsequent literature has raised compelling empirical objections. Stratal Optimality Theory (Kiparsky 2000, 2008) limits the number of cophonologies (layer types) in any given language to three, which are strictly ordered. The virtue of limiting strata in this way is that it draws attention to general properties of stems, words, and phrases, but often at the expense of being able to describe more "minor rules." Close studies of strata in agglutinating languages, for example, have generally resulted in the postulation of more than three levels below the word level alone (see e.g. Hargus (1985) on Sekani, Mohanan (1986) on Malayalam, Buckley (1994) on Kashaya). Both Hargus and Mohanan, like Czaykowska-Higgins (1993, for Moses-Columbian Salish) and Inkelas and Orgun (1998, for Turkish), argue in addition that the strata necessitated to describe the morphophonological subgeneralizations in the languages in question cannot be crucially ordered in the way that level ordering theory would require. It is important, however, to emphasize that the essence of level ordering theories is the same as the essence of cophonology theory, namely that the interleaving of

phonology and morphology is due to the association of morphological constructions with particular phonological patterns.

2.3 Paradigm Uniformity

A promising avenue of research on "optimal paradigms" seeks to examine whether paradigm-level considerations could motivate or even supplant cyclic cophonological models. This is especially promising in cases of recessive phonological alternations in which stem structure is preserved under subsequent affixation. The overall result is that paradigms are kept level, exhibiting, with phonologically uniform stem shape. It has been proposed that rather than resulting from cyclicity, stem uniformity effects follow from paradigm uniformity constraints which keep the shared portions of morphologically related words phonologically identical; see, for example, the Base-Identity constraints of for example, Kenstowicz (1996). When evaluated only with respect to the subconstituents of a single word, Base-Identity constraints function like high-ranked input-output faithfulness on a cophonology account, causing structure that is optimal for the innermost morphological constituent to persist even if outer layers of morphology render it phonologically opaque. This occurs in Turkish, as discussed earlier: lexically stressed roots (e.g. *lokánta* 'restaurant') and derived stressed stems (e.g. *süt-lű-ce* 'milk- ASSOC-MIT') keep their stress when they combine with would-be stress-assigning suffixes like pre-stressing predicative *-(i)di*; examples include *lokánta-ydı*, *süt-lű-ce-ydi.* The recessive character of stress-assigning suffixes can be attributed to paradigm uniformity: the derivational and inflectional paradigms of a lexically stressed noun like *lokánta* 'restaurant' all share an identically stressed root (*lokánta, lokánta-lar* (-PL), *lokánta-da* (-LOC), *lokánta-lar-da* (-PL-LOC), *lokánta-ydı* (-PRED), and so on.).

Of course, Base-Identity is not absolute in Turkish; it is only stressed roots whose phonological stress pattern is maintained across the paradigm. Lexically stressless roots alternate systematically, according to whether they combine with a stress-neutral suffix, for example, *araba-yá* 'car-DAT', or a stress-assigning suffix, for example, *arabá-yla* 'by/with car'. It is also important to note that the definition of "base" of a paradigm must be broadened to include not just roots but also complex stems. While the root *araba* is not inherently stressed, and therefore varies in shape depending on morphological context, a stressed stem like *arabá-yla* keeps its stress when suffixed, for example, *arabá-yla-m*$_1$ 'car-ASSOC-INTERROGATIVE = by/with car?'. Thus "base" is equivalent to "subconstituent" in a layering theory.

The predictions of paradigm constraints diverge from the predictions of cophonological layering models when applied to the shared stems of words neither of which is a subconstituent of the other. For example, Kenstowicz (2005) discusses the case of Spanish diminutives, formed by adding *-cito* [sito] (m.) /*-cita* [s*ita*] (f.) when the base ends in [n] or [r] and by adding *-ito/-ita* when the base ends in a vowel.[1] Examples cited by Kenstowicz, using his phonemic transcription, include [limon] 'lemon (m.)' → [limon-sito], [barko] 'ship (m.)' → [bark-ito], [korona]

'crown (f.)' → [koron-ita]. For nouns that have feminine and masculine gender counterparts, like [raton] 'mouse (m.)', [raton-a] 'mouse (f.)', the surface conditions for attachment of the [-sita/-sito] diminutive formatives are met by the *n*-final masculine but not by the *a*-final feminine. On Kenstowicz's assumption that the form of the non-diminutive noun determines the dimunitive suffix that is added, the feminine diminutive of 'mouse' should be [raton-ita], based on [ratona], whereas the masculine diminutive of 'mouse' should be [raton-sito], based on [ratón]. In fact, however, both diminutives have the diminutive formative triggered by an *n*-final input: [ratonsito], [ratonsita]. Kenstowicz proposes a paradigm uniformity analysis, which he attributes to Aguero-Batista, on which masculine and feminine diminutives are required to have the same surface stem shape. Masculine [ratón] transparently selects [-sito] ([raton-sito]), and by paradigm uniformity, the feminine [ratona] is required to select the [-sita] allomorph as well. Paradigm uniformity favors [ratonsita], while transparency of suffix selection favors [ratonita]; paradigm uniformity wins out. (There is an alternative to invoking paradigm uniformity in this case, namely treating gender-unspecified [ratón] as the input both to [raton-sit-o] and [raton-sit-a]. The argument for paradigm uniformity as a constraint is only as strong as the argument that nouns are gender-marked in the input to diminutivization. Since the diminutive endings themselves encode gender, this assumption could be questioned.)

A particularly interesting set of examples of paradigm uniformity is cited by Downing (2005a: 24, 130 ff.), in a study of suffix doubling in Jita (Bantu). In Jita verbs, the causative suffix *-y* triggers mutation (spirantization) of any preceding /r/: /gur-a/ 'buy-FV' → [gura], vs. /gur-y-a/ 'buy-CAUS-FV' → [gusya]. Jita has at least two other derivational suffixes with which the causative can co-occur: applicative /-ir/ and reciprocal /-an/. When the causative co-occurs with either of these, it must double, occurring both directly after the root and directly after the other suffix, for example, /gur-y-ir-y-a/ 'run-CAUS-APPL-CAUS-FV' → [gusi:sya] or /gur-y-an-y-a/ 'run-CAUS-RECIP-CAUS-FV' → [gusyanya]. In verbs with causative, reciprocal *and* applicative suffixes, the causative must occur three times: /gur-y-ir-y-an-y-a/ 'run-CAUS-APPL-CAUS-RECIP-CAUS-FV' → [gusi:sya:nya]. Similar multiplication of the causative occurs in Kinande (Mutaka and Hyman 1990) and Cibemba, among other Bantu languages (Hyman 1994, 2003). According to Downing, the multiplication of the Jita suffix occurs under pressure from paradigm uniformity. Downing proposes that the causative suffix is always the one added morphologically first to the root (thus, for a verb with all three suffixes, the abstract underlying structure is /Root-CAUS-APPL-RECIP-/. Phonologically, however, the causative /-y/ is always required to be last in the stem, by a right-alignment constraint. Crucially on Downing's analysis, the phonological form of the (always innermost) Root-CAUS subconstituent is required to be uniform across all causative forms of a given stem. The only way to satisfy both the uniformity and the right-alignment requirement is to add the causative more than once. In the applicativized causative /gur-y-ir-y-a/, for example, the /gur-y- . . ./ portion satisfies stem uniformity while the /. . . -y-a/ portion satisfies rightward y-alignment. Downing argues (p. 128) against an alternative cyclic account of causative doubling facts,

such as the one proposed for parallel affix doubling facts in Cibemba and a number of other Bantu langauges by Hyman (1994, 2003), on the grounds that there is no other evidence for cyclicity in Jita.

2.4 Paradigm Contrast

Another manifestation of paradigmatic considerations is the morphological (or lexical) need to keep words or stems phonologically distinct from one another; this need for paradigm contrast has been argued to inhibit or trigger phonological effects.

For example, Crosswhite (1999) argues on the basis of evidence in the Trigrad dialect of Bulgarian that an otherwise general rule of vowel reduction is blocked just in case it would cause the merger of two words in the same paradigm. In Trigrad Bulgarian, unstressed /o/ and /ɔ/ surface as [a], merging with underlying /a/:

(6)	/rog-ave/	['rogave]	'horns'
	/rog-ave-te/	[raga'vete]	'the horns'
	/sɔrp-ave/	['sɔrpave]	'sickles'
	/sɔrp-ave-te/	[sarpa'vete]	'the sickles'
		cf. [a'rala]	'plough'

Crosswhite observes that unstressed /o/ fails to reduce in a number of suffixes, for example, the *-o* ending on nominative masculine animate nouns: ['ago] 'older brother (nom.)', not *['aga]. Crosswhite observes that /o/ reduction fails precisely when, as in these cases, two distinct suffixes (one with /o/ and one with /a/) would merge if reduction applied. The accusative ending on masculine animate nouns is *-a*, as in ['aga] 'older brother (acc.)'. According to Crosswhite (and Kenstowicz 2005), vowel reduction is blocked when it would merge the nominative and accusative paradigm cells of masculine animate nouns. For a recent survey of these and other effects in which a neutralizing alternation is claimed to be blocked by a constraint against homophony, see Ichimura 2006.

According to Kurisu (2001), anti-homophony considerations can also trigger dissimilatory phonological alternations. Kurisu interprets a number of effects previously described as realizational morphology (see Section 5) as resulting from the requirement that input and output forms be distinct. On this view, process morphology is a repair of what would otherwise be the null realization of a morphological construction. Examples include the use of ablaut to mark plural in German (*Vater* ~ *Väter* 'father(s)', *Mutter* ~ *Mütter* 'mother(s)', p. 191), and the use of vowel deletion to derive deverbal nouns from infinitives in Icelandic (*klifra* 'climb-inf' → *klifr* 'climbing', *puukra* 'conceal (inf.)' → *puukr* 'concealment', p. 31, citing Orešnik 1978; Arnason 1980; Kiparsky 1984; Itô 1986; Benua 1995). Kurisu's analysis is that these constructions consist, morphologically, of zero-derivation, but that anti-homophony considerations compel the phonology to alter the output to avoid identity with the input. The fact that ablaut (in German) or vowel deletion (in Icelandic), are the preferred options, as opposed to any other imaginable

changes, follows, in Kurisu's account, from the ranking of faithfulness constraints penalizing deletion, insertion, and/or featural changes.

A challenge for Kurisu's view comes from cases of morphologically conditioned phonological effects applying alongside affixation, for example, German: *Gast* ~ *Gäst-e* 'guest(s)' or *Gaul* ~ *Gäul-e* 'pack horse(s)', with suffixation *and* ablaut (p. 191). Since affixation alone suffices to make two word-forms distinct in these cases, what motivates the accompanying ablaut effect? Kurisu's answer is that these cases are instances of double morphological exponence resulting from morphological opacity: the affixes in these examples are essentially invisible to the anti-homophony principle that requires the singular and plural cells of the paradigm to be distinct. The "first" layer of morphology is null, and phonology conspires to make the zero-marked plural stem (*Gäst*) distinct from the singular stem (*Gast*). The second layer of morphology, to which the phonology is blind, then double-marks the plural with a suffix: *Gäst-e*. Of course, double exponence is not limited to cases of this kind in which one exponent is arguably a phonological modification and the other is an overt affix; languages are known to use two or more overt affixes, or a suppletive stem plus overt affix(es), to mark a single category as well (e.g. Anderson 2001; Bobaljik 2000; Harris 2008a). Thus when ablaut is one of the two exponents of a morphological category, it could be analyzed, per Kurisu, as a phonological resolution to anti-homophony, or it could be attributed to whatever morphological factors are responsible for multiple exponence more generally.

Further afield, Wedel and Ussishkin (2002) have suggested that neutralizing phonological alternations can be inhibited if the words they would apply to exist in dense phonological lexical neighborhoods, that is, if there are high numbers of phonologically similar words in the lexicon. If this hypothesis is correct, contrast preservation might inhibit phonological alternations not only when the words in question are in the same paradigm, but even when they are morphologically unrelated. Dispersion might thus play an active role synchronically, not just the diachronic role suggested by Frisch, Pierrehumbert, and Broe (2004) in their discussion of Arabic root consonants. Frisch, Pierrehumbert, and Broe show that the distribution of consonants in Semitic roots is skewed to favor triples of root consonants that are phonologically internally disparate over triples of root consonants that are internally similar. Frisch, Pierrehumbert, and Broe suggest a diachronic path by which dissimilatory phonological pressures affect the lexicon. Whether the pressures are purely diachronic or also synchronic is a question that future research is sure to focus on.

Whatever the nature of contrast preservation principles turns out to be, the principles clearly play a subordinate role in grammars. Phonological alternations and neutralizations are rampant, as is the creation of homophony in paradigms. Even setting aside all cases of systematic syncretism within paradigms (see e.g. Baerman 2005), we still find numerous situations in which phonological neutralizations create homophony. To take just one example, in Russian the neutralization of unstressed /a/ and /o/ produces homophony between nominative/accusative and genitive forms of neuter *o*-stems (Baerman 2005: 809). A desinence-stressed

stem, for example, 'wine', has distinct nominative/accusative (*vin[ó]*) and genitive (*vin[á]*) forms, but a root-stressed stem, for example, 'place', is identical in both contexts (*mést[ə]*), due to vowel reduction.

2.5 Non-derived Environment Blocking (NDEB)

It has been widely observed that neutralizing phonological alternations which are triggered at morpheme boundaries fail to apply when the same phonological environment occurs morpheme-internally. "Derived environment effects," or "non-derived environment blocking" (NDEB), has been generally attributed to contrast preservation pressures, although formal accounts of the phenomenon vary widely. The classic example of a derived environment effect occurs in Finnish: as noted by Kiparsky (e.g. 1993b), the neutralizing assibilation alternation converting /t/ to /s/ before /i/ applies regularly at stem-suffix boundaries but does not affect morpheme-internal /ti/ sequences: *tilat-a* 'order-INFINITIVE' ~ *tilas-i* 'order-PAST', *but *silat-a, *silas-i.*

It was thought in the 1970s and 1980s that NDEB effects were associated with the class of cyclic, structure-changing rules; "Strict Cycle" principles proposed by Kiparsky (1982b) and Mascaró (1976) formalized this apparent correlation as part of the theory of Lexical Morphology and Phonology. However, subsequent findings (e.g. Hualde 1989a; Kiparsky 1993b) undermined the Strict Cyclicity correlation, showing that NDEB effects were not restricted to cyclic or to structure-changing rules and that not all cyclic or all structure-changing rules exhibit NDEB effects. A later wave of proposals, couched in Optimality Theory, focused on the tension between preserving input substrings from alteration, the idea being that morpheme-internal substrings (e.g. Finnish *ti*) would be preserved, but derived substrings (*t-i*) would be subject to alternation (e.g. Burzio 1997; Itô and Mester 1996b; McCarthy 2003a). A related approach is taken by Lubowicz (2002), who suggests that NDEB effects are those which apply only when input faithfulness has to be disrupted for some other reason, for example, resyllabification.

More recent work has gone back to the intuition that was first advanced by Kiparsky in the 1960s, namely that NDEB effects preserve contrast. Kiparsky's (1968a) Alternation Condition, though flawed in detail and later abandoned by Kiparsky (1982b) in favor of the Strict Cycle condition, captured the generalization that a given morpheme will undergo a neutralizing phonological alternation only if there is a contrast between morphological contexts in which the alternation is applicable to that morpheme and contexts in which the alternation is not applicable, making it possible for the underlying form of the morpheme to be recoverable by the learner.

For example, in Finnish, the initial *t* of *tilat* is always in the context of the Assibilation trigger *i*. By the Alternation Condition, it cannot alternate. The final *t* of *tilat*, however, sometimes occurs in an Assibilation context and sometimes does not. As a consequence it may alternate between *t* and *s* without obscuring the lexical contrast between stem-final /t/ and /s/.

The Alternation Condition has found recent new life in work by Lubowicz (2003), who proposes that neutralizing alternations be constrained by a grammatical pressure to preserve contrast. Morphemes that contrast underlyingly should not be neutralized in every possible surface context in which they might occur.

As with anti-homophony, it is not clear that the effects labeled by various analysts as NDEB are all of the same type, functionally or formally. Some effects which could be classified under NDEB are more likely due to the restriction of the pattern in question to a particular cophonology within the language. In Japanese, for example, a condition of bimoraic minimality is imposed on affixed words, leading to vowel lengthening and/or inhibiting the degree to which suffixed stems can be truncated; but the requirement is not imposed on bare roots, even when used as words (Itô 1990). A similar minimality phenomenon in Turkish is documented in Itô and Hankamer (1989) and Inkelas and Orgun (1995). Though the specifics of their analyses differ, these authors essentially characterize the minimal size restrictions as properties of stems of a particular morphogical type. Roots are not stems of this type, and evade the minimal size condition by virtue of its never being imposed on them at all. This same sort of analysis is given by Yu (2000) to the phenomenon in Tohono O'odham whereby final secondary stress is prohibited except in morphologically complex words. Yu provides an explicitly cophonological account in which the assignment of secondary stress to final syllables is part of the cophonology of word-formation constructions, but not part of the cophonology applied to roots, even those used as words.

Cophonological accounts such as these have little to say about local NDEB effects of the type seen in Finnish; conversely, accounts of local NDEB effects do not extend to the more global effects seen in Japanese, Turkish, and Tohono O'odham. The Alternation Condition, whether in its original form or in Lubowicz's more modern incarnation, is not applicable to Tohono O'odham secondary stress, which is not neutralizing.

It could well be that there are simply two types of NDEB effects, which cannot be merged: those involving neutralization, which are typically segmental and therefore typically local and for which a contrast preservation approach is appropriate; and those which involve prosody, which are not local and do not involve contrast neutralization, for which the cophonological accounts are suited (Inkelas 2000). The typology of NDEB effects is clearly an area of ongoing research.

2.6 *Locality and Bracket Erasure*

An important question for any model of the morphology-phonology interface is whether phonological patterns applying to one subconstituent of a word can make reference to properties of embedded structure.

The existence of NDEB effects suggests that phonology needs to distinguish complex from simplex structures. In Finnish it is necessary for the phonology, when applying to a form like /tilat-i/, to have access to the information that *tilat* and *-i* differ in their morphological status.

Beyond NDEB proper, many other interactions between morphology and phonology have been analyzed using phonological rules or constraints that directly reference morpheme boundaries or morpheme identity. One type of evidence that is frequently adduced is root prominence. In Turkish, for example, the presence of a lexically accented morpheme (root or affix) in a word overrides the default assignment of stress to the final syllable. However, when both a lexically accented root *and* a lexically accented affix combine in the same word, one must disappear, since Turkish words have only one stress each. In Turkish it is always affix stress which disappears, giving rise to the appearance of what McCarthy and Prince (1995) have characterized as a universal principle of root-faithfulness: grammar is always more faithful to root structure than to affix structure, in situations where it is necessary to choose. Alderete (1999, 2001) has analyzed similar root-prominence effects in Cupeño and Japanese, characterizing them in terms of root-faithfulness. It does not particularly matter, of course, whether the analytical tool is a root faithfulness constraint or something else; what matters is that the phonological grammar must be sensitive to the distinction between roots and affixes. In cases like these, some cophonological accounts have a different interpretation of what is going on. It is possible, in cophonology theory, to treat affixation like realizational morphology, the result of a morphologically-specific phonological mapping that takes a stem as input and produces an output that includes what in more traditional morpheme-based accounts would be called the affix (Orgun 1996; Inkelas 1998). On this implementation of cophonology theory, the phonological substance of the "affix" is not present in the input; only the phonological substance of the root is present. The asymmetry between root and affix on this account does not require reference to morpheme boundaries or to the identity of morpheme types; it only requires reference to input. The most extreme view of the relevance of morpheme boundaries to phonology, then, would be that the rules or constraints within a particular cophonology are completely insensitive to morphology, and that morphological sensivity arises only indirectly by means of the association of different cophonologies with different morphological word-building constructions. The most permissive view would grant phonology access to all kinds of morphological information. This was the original assumption in generative phonology (Chomsky and Halle 1968), in which all morpheme boundaries were visible to phonological rules, and is still prevalent in the Optimality Theory literature (e.g. McCarthy and Prince 1993).

A view that falls somewhere in between was developed in the 1980s in the general Lexical Morphology and Phonology framework (Kiparsky 1982b; Mohanan 1986), in which it was assumed that phonological rules applying on a particular cycle (or stratum, if non-cyclic) could see morpheme boundaries created on that cycle (or in that stratum), but that once such rules had applied, the internal morpheme boundaries would be "erased" or on some principle made invisible to rules applying on a subsequent cycle (stratum) of morphology. (A version of this principle of bracket erasure can be found in Chomsky and Halle (1968) as well.) The question of bracket erasure and the relevance of morpheme boundaries has drawn little direct attention since the rise of Optimality Theory, aside from a few

works such as Orgun and Inkelas 2002, Itô and Mester 2002; Shaw 2009). While many analyses in Optimality Theory allow phonological constraints to directly reference all embedded morphological structure, it is not always clear whether this follows from necessity or from convenience.

3 When Phonology Affects Morphology: Combinatorics

Thus far we have focused on cases in which phonological patterns differ across different morphological zones of complex words. In this section we examine a different kind of interface, in which word-formation possibilities can themselves be constrained by phonology, either because of phonological requirements on inputs to word formation or because of phonological requirements on the outputs of word formation. Constraints on word formation can result in the choice of one suppletive allomorph over another, or they can result in morphological gaps, where no output (or only a periphrastic output) is possible. There are even cases in which it appears that affix ordering is phonologically determined.

3.1 Suppletive Allomorphy

Thus far we have discussed interactions ascribable to grammar. Suppletive allomorphy is a type of morphology-phonology interface which involves the lexicon. Suppletive allomorphy is familiar to every beginning morphology student as the situation in which a given morphological category has two or more exponents which cannot be derived from a common phonological form but must be stored separately. Suppletive allomorphy enters the realm of the morphology-phonology interface when the choice between or among suppletive allomorphs is phonologically determined.

In a number of such cases, the distribution of suppletive allomorphs appears to resonate with phonological patterns in the language, suggesting that the phonological grammar could be responsible for handling the allomorphy. In Modern Western Armenian, for example, the definite article takes the shape *-n* following vowel-final nouns (e.g. *katu-n* 'cat-def') and *-ə* following consonant-final nouns (e.g. *hat-ə* 'piece'); Vaux 1998: 252.[2] Similar effects are familiar from Korean, in which several suffixes exhibit V- and C-initial suppletive allomorphs which occur after C- and V-final stems, respectively. Thus, the nominative, accusative, and topic-marked forms of *param* 'wind' are *param-i, param-il* and *param-in*, vs. the corresponding forms of *pori* 'barley': *pori-ka, pori-ril, pori-nin* (Paster 2006: 67, citing Odden 1993: 133). As researchers such as Mester (1994), Kager (1996), Anttila (1997a) and others have observed, constraints optimizing syllable structure (e.g. NoCoda, or Onset) would automatically entail the selection of allomorphs which produce CV syllables over those resulting in heterosyllabic consonant clusters (e.g. **hat-n*, in Armenian) or vowel sequences (e.g. **katu-ə*).

In a broad cross-linguistic survey of suppletive allomorphy, Paster (2006) uncovered a continuum of cases: some suppletive allomorphy (especially cases conditioned by syllable or metrical structure) is easy to characterize as phonologically optimizing, while other cases of allomorphy seem arbitrary or even non-optimizing. Consider, for example, the case of Haitian Creole, in which a particular determiner takes the form *-a* following vowels (e.g. *panié-a* 'the basket', *trou-a* 'the hole') and *-la* following consonants (e.g. *pitit-la* 'the child', *madãm-lã* 'the house') (Paster 2006: 86, citing Hall 1953: 32, via Klein 2003). This is the exact opposite distribution from the pattern just demonstrated in Korean, yet overall the syllable structures of the two languages are similar. If one allomorphic distribution makes sense phonologically, the other cannot. Or take Armenian noun pluralization: according to Vaux (1998: 31), "monosyllabic nouns take the suffix *-er* . . . and polysyllabic nouns take the suffix *-ner*: *ʧaʃ, ʧaʃ-er* 'meal(s)', *dodo, dodoʃ-ner* 'toad(s).' "

For apparently arbitrary phonologically conditioned allomorphy of this kind, lexical subcategorization is a common approach (e.g. Kiparsky 1982b; Inkelas 1990; Booij 2001; Paster 2006). The lexical entry includes all suppletive allomorphs, some or all of which are listed with a selectional frame identifying the phonological environment. In the case of the Armenian noun plural, for example, at least one of the suffix allomorphs must stipulate the number of syllables that the base of affixation is required to have. The other one can be the elsewhere case, if desired: {[[σ] er], [[ ] ner]}.

Paster argues for a subcategorization approach in all cases of suppletive allomorphy, even those, which the allomorph distribution could be attributed to grammar rather than the lexicon. Her argument is partly based on the fact that suppletive allomorphy is often opaque, conditioned by input factors which are obscured in the output by phonological alternations affecting the derived stem. In such cases, input conditioning is necessary even if the distribution of allomorphs makes phonological sense. Paster discusses the example of Turkish, in which the third-person possessive suffix has two suppletive allomorphs: *-I*, used after consonant-final stems (*ev-i* 'his/her/its house'), and *-sI*, used after vowel-final stems (*anne-si* 'his/her/its mother') (Lewis 1967; see also Paster 2006: 99). This distribution is rendered opaque when intervocalic velar deletion applies to a suffixed stem. In the third-person possessive, a velar-final word like *inek* 'cow' combines with the *-I* allomorph, as expected since *inek* is consonant-final. However, the result of velar deletion is [ine.i] (orthographic *ineği*), with the "wrong" surface allomorph. A surface optimization approach, given the choice between [ine.i] and [inek-si], would almost certainly be expected to pick [inek-si] (or even [ine-si]); CC clusters across morpheme boundaries, as would occur in [inek-si], are commonplace and never repaired by deletion or epenthesis, whereas VV clusters across morpheme boundaries are tolerated at no other stem-suffix junctures in the language. For these reasons, Paster analyzes this case not as output optimization but purely as input selection.

In one very interesting case of opaque allomorph selection in Polish, Lubowicz (2007) cites phonological contrast preservation as the motivation for the choice between suppletive allomorphs. The locative in Polish has two suppletive

allomorphs: *-e* and *-u*. Like other front suffix-initial vowels in Polish, *-e* triggers palatalization of a stem-final coronal consonant: *lis[t]* (nominative), *o liś[ć]-e* (locative) 'letter'. Exactly those stems whose final consonant is underlyingly palatal take *-u* instead: *liś[ć]* (nominative), *o li liś[ć]-u* (locative) 'leaf'. Lubowicz attributes the selection of the *-u* allomorph to contrast preservation. Exactly when *-e*, the preferred allomorph, would merge the contrast between underlyingly plain and underlying palatal root-final coronal consonants, *-u* is selected instead. It is important to note that the contrast being preserved here is an abstract phonological one. While "letter" and "leaf" form a minimal pair, the same distribution of *-e* and *-u* is found with roots that are independently distinct in other ways, for example, *łobu[z]* (nominative), *o łobu[ź]-e* (locative) 'troublemaker', but *pa[ź]* (nominative), *o pa[ź]-u* (locative) 'type of butterfly'.

3.2 Phonologically Motivated Morphological Gaps

In phonologically conditioned suppletive allomorphy, phonological grammatical constraints or the phonological requirements of individual affixes control which stems can combine with which affix allomorphs. Sometimes the phonological grammar, or the phonological selectional requirements of individual affixes, can be so strict as to block morphological combination altogether, resulting in phonologically driven morphological gaps. For many speakers of Turkish, suffixation is grammatical only if the resulting word is at least disyllabic (Itô and Hankamer 1989; Inkelas and Orgun 1995). In Dutch, the superlative ending *-st* cannot be added to adjectives ending in [is], [sk], [st]; thus *bruusk* 'sudden' has no lexical superlative counterpart (**bruusk-st* [bryskst]) but must enter into a periphrastic syntactic alternative: *meest bruusk* (Booij 2005). In Tagalog, infixation of the agentive focus marker *-um-* is impossible if the stem begins with /m/ or /w/, creating paradigm gaps for such words (Schachter and Otanes 1972; Orgun and Sprouse 1999). A number of similar cases are surveyed by Carstairs-McCarthy (1998). Phonologically conditioned gaps differ from suppletive allomorphy in that there is no "elsewhere" allomorph; without this alternative, the word simply cannot be formed, resulting in a gap. Often there is a syntactic alternative; in English, for example, the comparative suffix *-er* attaches only to (loosely speaking) monosyllabic stems; thus *vast-er* but **gigantic-er*, forcing speakers to resort to the periphrastic *more* comparative: *more gigantic* (Poser 1992).

3.3 Haplology Effects

Menn and McWhinney (1984) draw attention to a common cross-linguistic pattern of prohibiting sequences of homophonous morphemes, which they term the Repeated Morph Constraint (RMC). A well-known example occurs in English, where the possessive ending /z/ is not added to – or at least not realized on – words ending in the homophonous plural suffix /z/; thus *dogs* and *dogs'* are pronounced identically ([dagz]) in the phrases *the dogs hate their collars* and *the*

dogs' collars drive them crazy. Irregular plurals take the possessive (*children's*) and so do words ending in the strings homophonous with allomorphs of the plural (i.e. [(ɨ)z] or [s]), for example, *Katz's* [kætsɨz] or *cats* [kæts]. The RMC may, according to Menn and McWhinney, result in in haplology, as in the English example of *cats'*, where a single phonological exponent [s] stands for what appear to be two morphemes. In other cases the RMC can also trigger suppletive allomorphy or avoidance, in which a periphrastic alternative is preferred. For example, the English adverbial *-ly* ending (*quick* (adj.), *quickly* (adv.)) cannot combine with those adjectives already ending in *-ly,* for example, *manly* or *heavenly*: **manlily, *heavenlily* (adv.). The lexical gap which a word like **manly* cannot fill must be approximated by a phrase like *in an manly fashion*. Since it affects word form and creates morphological paradigm gaps, the RMC generalization would seem to be a clear case of phonology interfering with morphology.

The RMC is, however, clearly not universal, even within a language. There exist many unperturbed sequences of homophonous morphs; there are also instances of suppletion and avoidance in which morph repetition is not an issue. English, for example, permits sequences of the plural or possessive followed by the homophonous reduced form of *is,* for example, *one of the cats's* [kætsɨz] *trapped in the closet!* or *Whose guacamole do you like best? John's is* [dʒɑnzɨz] *clearly the winner*. The RMC applies only to sequences of plural and possessive, not to all sequences of /z/ morphemes. Another example occurs in Turkish, which uses the same suffix (/-I/~/-sI/) both to mark third-person singular possessors (*aile* 'family', *aile-si* 'his/her/its family'; *araba* 'car', *araba-sı* 'his/her/its car') and also as a marker at the end of head-modifier compounds, in which the possession relation, if any, is quite abstract (Lewis 1967: 42): *aile araba-sı* 'family car'. The possessive suffix cannot occur twice in succession; therefore, in isolation, compounds like *aile araba-sı* are actually ambiguous between a possessed (e.g. 'his/her/its family car') and unpossessed reading (Lewis 1967: 46). The ungrammaticality of a doubly affixed possessive compound, such as **aile araba-sı-sı*, cannot, however, simply be attributed to the RMC. Other possessive suffixes, for example, first person possessive /-m/, "associative" -/lI/ and "occupational" /-CI/, are also in complementary distribution with the compound-marking possessive suffix even though they are not homophonous with it (Lewis 1967: 49–50): *aile araba-m* 'my family car', not **aile araba-sı-m*, etc. Thus even this apparently transparent case of a repeated morph prohibition on possessive /-I/~/-sI/ turns out to be part of a more general pattern of morpheme co-occurrence. How, then, are we to know whether affix co-occurrence restrictions between homophonous affixes are a distinct subtype of affix co-occurrence restrictions generally? Further research is required in this area, but answers are likely to be of two types. One is statistical: if homophonous affix pairs form a larger than expected subset of the class of morpheme pairs that cannot occur next to each other, the RMC would be supported, though this task would be hard to accomplish given current data. A second possible answer would be to show that the lexical gaps or lexical ambiguities resulting from RMC effects pattern differently from those resulting from other, more arbitrary morpheme co-occurrence restrictions.

3.4 Linear Order

A number of cases have been described in which phonology constrains the linear order of morphemes. Mortensen (2006) has collected a very interesting set of examples in which constituents in coordinate compounding are ordered according to their phonological properties, principally vowel quality and tone. In one dramatic case from Jingpho, which Mortensen draws from a 1990 monograph in Chinese by Qingzia Dai, the order of elements in compounds with coordinate semantics follows from the height of the tonic (root) vowels: the stem with the higher vowel always precedes the stem with the lower vowel. Thus *lùʔ-ʃá* 'drink-eat = food' is a grammatical compound, while **ʃá-lùʔ*, with the same presumed meaning, would be ungrammatical (Mortensen 2006: 222–223). Mortensen documents many such compounding cases, mainly involving vowel quality and/or tone, in which the order of elements follows a scale. Sometimes the scale is phonetically transparent, as in the Jingpho case of vowel height, and sometimes not, when historical changes have obscured the original phonetic or phonological basis for the scale.

Another area in which phonology determines linear order is found with "mobile affixes," discussed by Fulmer (1991), Noyer (1994), and Kim (2008). These vary freely between prefixal and suffixal attachment, with phonological considerations being the deciding factors. In the San Francisco del Mar dialect of Huave, for example, the subordinate marker *m* attaches as a prefix to vowel-initial bases (*m-[u-ty]* 'SB-TV-eat = (that) s/he eats') but as a suffix to consonant-initial bases (*[mojk-o]-m* 'face.down-V-SB = (that) s/he lies face down'). Similar behavior is exhibited by other affixes, including the stative *n: n-[a-kants]* 'ST-TV-red = red' vs. *[pal-a]-n* 'close-V-ST = closed' (Kim 2008: 332).

As proposed by Kim (2008) and, for similar facts in San Mateo Huave, by Noyer (1994), such cases can be modeled in Optimality Theory by the general schema proposed by McCarthy and Prince (1994a) in which phonological considerations ("P") outrank morphological considerations, such as affix ordering ("M"). In Huave, according to Kim, mobile affixes are preferentially suffixing (the "M" condition), but will prefix if suffixation would produce consonant clusters that would require epenthesis (the "P" condition) (pp. 340–341). Thus for a base like [a-rang] 'TV-do', *m-* prefixation (*m-a-rang*) is preferred over *m*-suffixation (**a-rang-m*, **a-rang-am*), since the latter would produce an unsyllabifiable cluster requiring repair. In cases where both prefixation and suffixation options would require epenthesis, suffixation is preferred: first-person *s* combines as a suffix with base *t-a-rang* 'CP-TV-do = did (it)' to yield *t-a-rang-as*, with epenthesis, rather than as a prefix (**s-tarang* or **sa-tarang*) (pp. 340, 342).

In general, however, the effect that phonology has been argued to play in the ordering of morphological elements is fairly limited. The great bulk of affix ordering is determined by the morphology, not by the phonology. To take a very simple example from Turkish, consider the interaction of the "occupational" suffix /-CI/ and case endings, such as the dative /-E/. Both can attach to roots. Turkish epenthesizes vowels to break up triconsonantal clusters, and epenthesizes

glides to break up vowel-vowel sequences at stem-suffix boundaries. Thus we find alternations like these: /jeni-CI/ → [jenidʒi] 'new-PROF' (*yenici*), /film-CI/ → [filɪ̯mdʒi] 'film-PROF = film-maker' (*filimci*); /jeni-E/ → [jenɪ̯je] 'new-DAT' (*yeniye*), /film-E/ → [filme] 'film-DAT'. When both /-CI/ and a case suffix occur in the same word, affix order is fixed. /-CI/, as a derivational suffix, always precedes case: /jeni-CI-E/ → [jenidʒije] 'new-PROF-DAT' (*yeniciye*); /film-CI-E/ → [fil**i**mdʒi**j**e] 'film-maker (dative)' (*filimciye*). In the latter example, two epenthesis operations are required to bring the syllable structure of the resulting word into conformity with Turkish requirements. By contrast, the alternative affix ordering would produce perfectly well-formed syllables with no need for epenthesis: /film-E-CI/ → [fil.me.dʒi] (**filmeci*). But **filmeci* is completely impossible in Turkish; phonological considerations do not trump morphological constraints on relative affix order.

Paster (2005) explores one well-known apparent case, from the Fuuta Tooro dialect of Pulaar (Fula), in which phonology has been claimed to order affixes. In a study of the Gombe dialect, Arnott (1970) observed that a number of C or CV suffixes in the same general "zone" of the word appear to occur in a phonologically determined order: all suffixes with "t" precede all suffixes with "d," which precede all suffixes with "n," which precede all suffixes with "r." Paster cites similar examples of this "TDNR" template from Fuuta Tooro Pulaar, for example, *jaɓ-t-id-ir-an-ii* 'take-INTENSIVE-COMPREHENSIVE-MODAL-DATIVE-PAST' and *yam-d-it-in-ir-ii* 'healthy-DENOMINATIVE-REPETITIVE-CAUSATIVE-MODAL-PAST' (Paster 2005: 164). As Paster argues, however, the order of affixes in both dialects of Pulaar conforms to semantic ordering principles of the kind articulated by Bybee (1985) and Rice (2000); there is no case in which the phonological TDNR template contravenes an ordering that one might otherwise expect on morphological grounds, and thus no clear evidence that phonology is interfering with morphology. Paster also observes that the TDNR template, in which consonant sonority increases from left to right, is not completely convincing as a phonological phenomenon, since in actual Pulaar words, vowels typically separate the consonants which correspond to the elements of the TDNR template. A sonority-based template like TDNR would make more sense, Paster argues, if consonants were being ordered by sonority in order to fit into a single syllable onset or coda, but this is not the case in Pulaar. In sum, Paster concludes, the Pulaar pattern is significant for coming closer than any other example to being a case of phonologically-driven affix sequencing but still not fully meeting that description.

The most substantial influence of phonological considerations on the linearization of morphemes is found in infixation, which is generally viewed as being just like affixation except that the affix is phonologically positioned within the stem instead of peripheral to it (see e.g. Moravcik 1977, 2000; McCarthy and Prince 1993; Yu 2007). The interest of infixation for the phonology-morphology interface lies in phonological generalizations about where in a word an infix can appear and about what, if anything, motivates infixation synchronically.

Surveys of infixation from Moravcsik (1977) to Yu (2007) have found a small and principled set of recurring sites for segmental infixes: next to the initial or final

consonant or vowel, or (in lexical stress languages) next to a metrical prominence. As Yu observes, these sites are defined in terms of types of elements that all words in the relevant language contain. All words contain consonants, vowels, and (in stress languages) stress. By contrast, there is apparently no evidence of infixation to syllables with particular tones, or to syllables containing particular types of segments (e.g. fricatives or ejectives) or even to heavy syllables (e.g. those with long vowels or consonant clusters). These are elements that languages do not typically require all words to possess.

Only dependent morphemes (affixes and, rarely, clitics; Harris 2000) have ever reliably been shown to infix; infixation is apparently not a possible property of compounding or phrasal combination, aside from the suggestive example of expletive infixation before a stress foot in English words (*amálgamàted* → *amálga-bloody-màted*, *Kàlamazóo* → *Kàlama-fuckin-zóo*, and so on McCarthy 1982).

A stimulating theory of infixation was introduced in the early 1990s, within the framework of Optimality Theory, by McCarthy and Prince (1993), who observed that many cases of infixation could be interpreted as improving the prosodic structure of the derived word in comparison to the structure the word would have if the infixed element were instead adfixed. Infixation was a key motivator in McCarthy and Prince's proposal that at least some phonological constraints "P" can outrank morphological constraints "M," particularly those having to do with the edge-alignment of affixes.

The most convincing examples brought forth for this view, termed the "Phonological Readjustment" view in Yu (2007), involve syllable structure. McCarthy and Prince's original example concerns the agentive focus marker, *-um-*, in Tagalog, which precedes the initial vowel: *bilih* 'buy' → *b-um-ilih*, *gradwet* 'graduate' → *gr-um-adwet*, and so on. According to McCarthy and Prince, *-um-* is a prefix, subject to the "M" constraint Align-L(um-, Stem), which is outranked by a "P" constraint banning closed syllables (NoCoda). In considering the possible locations for *-um-* in case of *gradwet*, the prefixing candidate *um.grad.wet* has two closed syllables, while the infixing candidate *gr-u.m-ad.wet* has only one closed syllable, satisifying NoCoda better. Infixation is preferred because of the P >> M ranking.

By contrast to numerous examples in which infixation can be interpreted as improving syllable structure, very few examples have been found in which infixation can be construed as improving segment structure (see e.g. Yu 2007: Chapter 6). This asymmetry raises doubts about the generality of the potential for "P" constraints to outrank morphological alignment.

A second major issue confronted by the P >> M model is locality. If infixation is misalignment with the aim of avoiding bad structures, the P >> M model predicts what Yu calls "hyperinfixation" or unbounded infixation, a point made also by Orgun and Sprouse (1999) and McCarthy (2003b). In Section 3.2 we mentioned that the Tagalog *-um-* infix is prohibited from combining with *m-* or *w-* initial stems. Modeling this prohibition as a phonological constraint (e.g. OCP, following Orgun and Sprouse), a P >> M ranking would predict that *-um-* could infix further into the word than usual to avoid the undesired *mu* or *wu* sequence. This does

not happen in Tagalog, nor does the comparable situation appear to arise in any other language: infixes are tightly restricted to appear near edges. McCarthy (2003b) addresses this problem for P >> M by modifying alignment constraints so that they are categorical; Tagalog *um* is subject both to a violable constraint forbidding it from being separated by a segment from the beginning of the word and to an inviolable constraint forbidding it from being separated by a syllable (or more) from the beginning of the word (p. 95 ff.).

A radically different approach to infixation is offered by Yu (2007), who observes that many cases of infixation fall outside the P >> M model in the sense of not being prosodically improving in any discernible way, yet still conform to the locality generalizations that were potentially problematic for the P >> M model. Yu points out that some cases of infixation neither improve nor worsen syllable structure For example, the Hua negative infix *-ʔa-*, which is CV in shape, infixes before the final syllable: *harupo* → *haru-ʔa-po* '(not) slip', *zgavo* → *zga-ʔa-vo* '(not) embrace', even though adfixing (*ʔa-harupo, harupo-ʔa*) would have produced equally good syllables (Yu 2007: 30, citing Haiman 1980). Other cases of infixation arguably make syllable structure worse. For example, the nominalizing *-ni-* infix in Leti follows the first consonant (e.g. *kaati* 'to carve' → *k-ni-aati* 'carving', *polu* 'to call' → *p-ni-olu* 'act of calling, call'), producing marked consonant clusters and vowel sequences that would be avoided by simple adfixation (e.g. *ni-polu, polu-ni*) or infixation to a different position (e.g. *po-ni-lu*) (Yu 2007: 28, citing Blevins 1999).

Yu concludes that locality and generality, rather than phonological optimization, are the main generalizations that a synchronic model of infixation should capture, and proposes a lexical subcategorization approach building on, for example, Broselow and McCarthy (1983), McCarthy and Prince (1986), Inkelas (1990). On this approach, each infix is associated with a lexical statement defining its position relative to one or both edges of the stem it combines with. Phonological entities to which such statements are permitted to refer come from a small list of "pivots" that cross-linguistically are shown to separate infixes from stem edges: segments, syllables, and stressed constituents (Yu 2007: 52).

4 When Phonology Affects Morphology: Form

Some morphological constructions are phonologically compositional, in the sense that the morphology combines two or more elements with fixed phonological shapes and the "regular" rules of the phonology apply to give the combination its surface phonological form, which varies with the shapes of the input morphemes. For example, prefixation of *pre-* in English (*pre-register*, *pre-ordained*) is a simple matter of concatenating the fixed string [pri] with a base. But in some morphological constructions, extrinsic considerations constrain or determine output phonological shape, with input morphemes conforming to output shape requirements instead of determining output shape themselves. We will survey two such phenomena here: templatic morphology and reduplication. These are commonly termed "prosodic morphology," because in each case a morpheme is

expressed phonologically in a way that is not constant across the set of stems formed from that construction but is predictable from construction-specific metrical or syllabic constraints on the phonological shape of the complex stem. Prosodic morphology is often described as a trio, with infixation as the third member; however, the considerations driving infixation are rather different, as seen above.

4.1 Templates

Templates are morphological constructions, typically associated with specific derivational or inflectional morphological categories, which directly constrain the phonological shape of the derived stem. McCarthy (1979a, 1981) broke new ground by analyzing the fixed shape of specific derivational subtypes of Arabic verbs as composed of templates consisting of CV timing units. These templates, each expressing a specific morphological category, combine with other morphemes which consist of consonants, and with others consisting of vowels, to form complex words. For example, the consonantal root /ktb/ 'write' combines with the "perfective passive" vocalic morpheme /ui/ and the "causative" template CVC-CVC to form *kuttib.* In their seminal 1986 paper, McCarthy and Prince showed the role of prosodic structure in defining the various templates not just in the Arabic root and pattern morphological system but more generally in prosodic morphology cross-linguistically. According to McCarthy and Prince, templates are always defined in terms of the universally accessible units of mora, syllable, and foot, rather than in terms of the C, V, or X timing units proposed in earlier work by McCarthy (1979a, 1981), Leben (1980), Hyman (1985), and others.

Sometimes templates constrain the shape of stems or words without contributing any particular semantic or syntactic function of their own. A simple case of this occurs with minimal word size constraints, which can compel epenthesis or other phonological augmentation strategies in short words. In Swati, as in many other Bantu languages, a disyllabic minimality requirement on words compels the use of a dummy suffix *-ni* in verbs that would otherwise be monosyllabic, a situation which arises in imperatives, the one morphological environment with no prefixes. Thus, while the infinitive of the stem /dlá/ 'eat' is *kû-dlá*, disyllabic by virtue of containing the infinitive prefix, the unprefixed singular imperative is *dlá-ni*, with augmentation that is not required for verbs formed from longer stems, such as /bóna/ 'see': *kû-bóna* 'INF-see = to see', *bóna* 'see (singular imperative)' (Downing 2006: 3). In Lardil, uninflected nouns are subject to apocope, seen in alternations like *wiwala-n* 'bush mango-NONFUT.ACC', *wiwala-ṛ* 'bush mango-FUT.ACC', but *wiwal* 'bush mango, from /wiwala/; *karikari-n* 'butterfish-NONFUT.ACC', *karikari-wuṛ* 'butterfish-FUT.ACC', but *karikar* 'butterfish' (from /karikari/) (Hale 1973: 424). Apocope is blocked when the result would have only one short vowel, for example, *kela-n* 'beach-NONFUT.ACC', *kela-ṛ* 'beach-FUT.ACC', *kela (*kel)* 'beach', from /kela/ (Hale 1973: 421). Uninflected nouns with only one short vowel are even subject to augmentation, so that they achieve bimoraic size: *ṭer-in* 'thigh-NONFUT.ACC', *ṭer-uṛ* 'thigh-FUT.ACC', but *ṭera* 'thigh', from /ṭer/ (Hale 1973: 427). Uninflected nouns in Lardil are clearly affected by the prosodic limitation

on word size. (For further discussion of Lardil, see, for example, Kenstowicz and Kisseberth 1979, Itô 1986, Blevins 1997, and Bye 2006).

A more involved case, in which prosodic templates constrain stem shape, occurs in Yawelmani (Archangeli 1983, 1991, based on Newman 1944). In the verbal system, each root and affix is lexically associated with one of three prosodic templates: a light syllable, a heavy syllable, and an iambic foot consisting of a light syllable followed by a heavy syllable. These templates determine the form of the root. When, in the same word, an affix and root are associated with conflicting templates, the one associated with the affix prevails, leading to root alternations in related stems. For example, the root "walk" has a default iambic template, as in *hiweet-en* 'will walk', but shortens to a heavy syllable and to a light syllable when combining with suffixes associated with the corresponding templates, as in *hewtihni* 'one who is roaming' (< *hiiwt-(ʔ)ihni*) and *hiwtiṅay* 'while walking' (< *hiwt-(ʔ)iṅay*) (Archangeli 1991: 232, citing Newman 1944: 101, 110, 136).

Prosodic templates can even specify segmental content. In Tiene verbs, derivational stems are constrained by a CV(C)VC- template whose consonants are subject to two major restrictions (Hyman and Inkelas 1997 and Hyman 2006a, based on Ellington 1977). In CVCVC- stems, the middle consonant (C_{med}) must be coronal and the final consonant (C_{fin}) must be non-coronal, that is, labial or velar. These restrictions can force the choice of infixal allomorphs of suffixes such as the stative, which has both an infixal allomorph (with coronal /l/) and a suffixal allomorph (with non-coronal /k/): *kab-* 'be divided' ~ *ka-la-b-* (stative); *yat-* 'be split' ~ *yat-ak-* (stative); *sɔn-* *'write'* ~ *sɔn-ɔŋ-* (stative). C2 and C3 must also agree in nasality, leading to nasal~oral alternations: *vwuɲ-* 'be mixed' + *-ek-* (stative) → *vwuɲ-eŋ-*; *dim-* 'become extinguished' + *-se-* (causative) → *di-se-b-*.

In recent work in Optimality Theory, starting with McCarthy and Prince 1994a, researchers have argued that templates are emergent artifacts of constraint interaction, rather than abstract structures manipulated by grammar. In Arabic and Tiene, for example, it might be possible, instead of stipulating baldly that the (derivational) verb stem must be CV(C)VC- in shape, or even simply that it must be bimoraic, to let this profile emerge from constraints like Ft-Bin (feet are binary) and All-Ft-Left (every foot must be initial), ranked high in the morphological environment of the verb stem. The emergent template approach has been applied fruitfully to many cases of reduplication by McCarthy and Prince (1995) as well as Gafos (1998a) and Hendricks (2001), among others, and has been extended beyond reduplication by Downing (2006). The motivation for deriving rather than stipulating templates is two-fold: first, deriving templates from independently needed markedness constraints should yield a more limited, principled set of possible templates than what it is possible to stipulate, and second, templates constrain form in ways other than simple prosodic size. Markedness constraints can constrain segmental form as well. The flexibility of emergent templates is useful in characterizing cases like Tiene, in which the restriction about consonantal place of articulation in verbs cannot be expressed by annotating particular prosodic positions for segmental features. C_{fin} is unrestricted in CVVC stems; C_{fin} is constrained, by dissimilatory principles, only if C_{med} is present. This kind of contingent restriction is better suited to constraints of the sort posited by Hyman

and Inkelas, in which C_{med} and C_{fin} must differ in place of articulation (and C_{med} must be coronal).

4.2 Reduplication

Reduplication is the doubling of some part of a morphological constituent (root, stem, word) for some morphological purpose. Total reduplication duplicates the entire constituent. It is often, though nowhere near always, semantically iconic, as in the duplication of nouns with human reference to form plurals in Warlpiri (*kurdu* 'child', *kurdu-kurdu* 'children'; *wirriya* 'boy', *wirriya-wirriya* 'boys') (Nash 1986: 130).

Partial reduplication, which exhibits a very wide range of meanings, usually involves a prosodically characterized template for the reduplicant. For example, McCarthy and Prince (1986) analyze the reduplicating prefix marking progressive aspect in Mokilese as a bimoraic syllable: *poadok* [pɔdok] 'to plant' → *poad-poadok* [pɔdpɔdok] ~ *poah-poadok* [pɔɔpɔdok] 'to be planting'; *piload* [pilɔd] 'to pick breadfruit' → *pil-piload* [pilpilɔd] 'to be picking breadfruit'; *kohkoa* [kookɔ] 'to grind coconut' → *koh-kohkoa* [kookookɔ] 'to be grinding coconut' (Harrison and Albert 1976: 60, 220). Typically the base exceeds the reduplicant template in size and thus the resulting reduplication is, as in these Mokilese data cited, partial. Occasionally, a reduplicant template will be bigger than the base; in which case reduplicant augmentation occurs. For example, the Mokilese form *pa* 'weave', which is monomoraic, reduplicates as *pah-pa* [paapa], with a bimoraic reduplicant (Harrison and Albert 1976: 60).

While typical reduplicant shapes are described in the prosodic units of mora, syllable, and foot, Moravcsik (1977) is credited with the observation that reduplication rarely unambiguously copies an *existing* mora, syllable, or foot from the base. Rather, as modeled by theories like Prosodic Morphology (McCarthy and Prince 1986) and, subsequently, approaches to reduplication within Optimality Theory (McCarthy and Prince 1994a, 1995), templatic requirements seem to be output requirements on the reduplicant. In Mokilese, what copies is enough base material to flesh out a heavy syllable reduplicant, even if the corresponding string is not itself a syllable in the base, (e.g. reduplicant [pil], from base [pilɔd]). In Optimality Theory this output orientation can be modeled by stating reduplicant shape as the output requirement RED = $\sigma_{\mu\mu}$. McCarthy and Prince (1994a) and Urbanczyk (1996), working in Generalized Template Theory, and, from a different angle, Downing (2006), have pursued the goal of deriving templates rather than stipulating them. In different ways, these researchers propose that reduplicants assume the canonical phonological form of whatever morphological constituent type (affix, stem, morphologically complex stem) they instantiate. This form does not have to be stipulated specially for the reduplicant but is motivated more generally for the language, or even cross-linguistically. Recent literature has suggested that some reduplication, particularly when limited to consonants, may not have even an indirect prosodic templatic characterization at all. Hendricks (1999, 2001) points to cases such as expressive reduplication in Semai, which copies the first and last consonant of the base: *pŋ-payaŋ* 'appearance of large stomach constantly bulging out', *cw-cruha:w* 'sound of waterfall, monsoon rain' (Diffloth 1976).

Partial reduplication usually duplicates that edge of the stem to which the reduplicant is closest, but opposite-edge reduplication (not including the dual-edge version found in Semai) occurs as well. A dozen or so cases are documented in surveys by Nelson (2003, 2004); Kennedy (2003); and Riggle (2003); all target the beginning portion of a base, for example Koryak CVC reduplication marking absolute case: *mɪtqa* 'oil' → *mɪtqa-mɪt; qanga* 'fire' → *qanga-qan.*

Partial reduplication is also commonly infixing, as in Chamorro (Topping 1973: 183), where habitual/continuative CV reduplication targets stressed syllables (*hátsa* 'lift' → *há-ha-tsa*, *hugándo* 'play' → *hugá-ga-ndo*) and intensifying CV reduplication targets the final syllable (*métgot* 'strong' → *métgo-go-t* 'very strong', *ñálang* 'hungry' → *ñála-la-ng)*; see, for example, Broselow and McCarthy 1983: 55–56). Internal reduplication usually duplicates adjacent material, as in these examples, but there are some exceptions to this. In Washo, for example, plural reduplication infixes a mora in the vicinity of the stressed syllable. In case the stressed syllable is closed, as in *nén.t'uš* 'old woman (nom.)' or *ʔéw.ʃiʔ* 'father's brother', reduplication copies a non-adjacent CV string: *ne.t'ún.t'uš-u* 'old women (nom.)', *ʔe.ʃíw.ʃiʔ* 'father's brothers' (Yu 2005: 440, citing Jacobsen 1964). Creek plurals are formed by infixing a copy of the stem-*initial* CV before the stem-*final* consonant (Riggle 2003, citing Booker 1980; Haas 1977, and Martin and Mauldin 2000): *holwak-íː* 'ugly, naughty' → *holwaː-ho-k-íː*, *falápk-iː* 'crooked' → *falap-fa-k-íː.*

4.2.1 Identity Effects in Reduplication: Over and Underapplication Since Wilbur's influential (1973) dissertation, researchers have paid special attention to phonological opacity arising in reduplication constructions. For example, consider Javanese total reduplication, which has pluralizing semantics and can apply to verbs and adjectives. When suffixed, for example, by demonstrative *-e,* reduplicated forms exhibit "overapplication" of intervocalic *h*-deletion and underapplication of closed syllable laxing and stem-final consonant devoicing (Inkelas and Zoll 2005: 146, 148, citing Dudas 1976: 207–208):

(7)

Gloss	*Stem*	*-demonstrative*	
'broken'	bəḍah	bəḍa-e	normal application of intervocalic h-deletion
	bəḍah-bəḍah	bəḍa-bəḍa-e	"overapplication" of h-deletion in first stem copy (demonstrative)
'cylindrical'	gilɪk	gilig-e	normal application of intervocalic voicing
	gilɪk-gilɪk	gilig-gilig-e	"underapplication" of stem-final devoicing in first stem copy (demonstrative)

Although opacity occurs outside of reduplication as well, its appearance in reduplicative examples like these is interpreted by Wilbur (1973), who posits a reduplicative Identity Principle, and in Base-Reduplicant Correspondence Theory, which posits Base-Reduplicant faithfulness constraints, as driven by the functional need to keep the two parts in reduplication – base and copy – segmentally identical. This imperative, while apparently obeyed in Javanese, is not satisfied in every case. For example, Urbanczyk (1996) and Struijke (2000) draw attention to a reduplicative construction in Lushootseed, illustrated by examples like *wális* 'type of frog' → *wá-w'lis* 'little frog', *caq'(a)* 'spear' → *ca-cq'* 'act of spearing big game on water', and so on. (Urbanczyk 1996: 167). Urbanczyk and Struijke analyze this pattern as CV prefixing reduplication accompanied by syncope in the base. On this account, underapplication of syncope would better preserve reduplicant-base identity (*wá-walis*); however, syncope applies transparently anyway without impedance from base-reduplicant identity constraints.

Some cases of reduplicative opacity can be attributed to layering or stratal aspects of the phonology-morphology interaction. With regard to the Javanese case above, Inkelas and Zoll (2005: Chapter 5) argue that demonstrative suffixation occurs prior to reduplication, triggering *h*-deletion and consonant voicing and preventing closed-syllable laxing from occurring; the suffixed stem (for example, *bəḍa-e* or *gilig-e*) is then input to reduplication, which copies the root as is, preserving the effects of the stem-level phonological alternations. What portion of reduplicative opacity will yield to layering accounts as proposed by Inkelas and Zoll (2005) and Kiparsky (2010), and what portion requires identity principles, is still an open question. It may be important, in deciding this question, to factor apart morphologically driven reduplication, such as the Javanese example, from phonologically-driven segment duplication. The latter clearly requires phonological identity principles (copying or correspondence, as appropriate to the theoretical framework in use). See Hendricks (1999, 2001), Yu (2005), Riggle (2006), Inkelas (2008a), and Pulleyblank (2009) for discussion relevant to the distinction between morphological reduplication and phonological copying.

4.2.2 Fixed Segmentism in Reduplication It is often the case that one of the two copies in morphological reduplication contains some fixed material which either co-occurs with or supplants material that would otherwise be expected to copy. An example of the former occurs in Khasi, where iterative verb reduplication connects the two copies of the verb with the linker *ši*, for example, *iaid-ši-iaid* 'to go on walking', *leh-ši-leh* 'keep repeating', *kren-ši-kren* 'keep talking' (Abbi 1991: 130, cited in Inkelas and Zoll 2005: 36). An example of the latter occurs in English, where an ironic/derisive total reduplication construction assigns "shm" to be the onset of the second copy, replacing an existing onset, if any: *fancy-shmancy*, *handsome-shmandsome*, *OT-shmOT,* and so on; see for example, Alderete *et al.* (1999). This phenomenon has been termed "Melodic Overwriting" (McCarthy and Prince 1986; Yip 1992; Alderete *et al.* 1999).

One functional motivation that has been offered by Yip (1997) for Melodic Overwriting is that it makes the two copies different. Support for this interpretation

is found in (a) the predominance of Melodic Overwriting in total, as opposed to partial, reduplication and (b) the fact that constructions involving Melodic Overwriting sometimes block when the two copies would be identical phonologically, or, perhaps more commonly, exhibit dissimilatory allomorphy which guarantees that the copy with the fixed material is different from the intact copy. In Turkish, as described by Lewis (1967: 237–238), a construction meaning "and so on, and suchlike" doubles a word and imposes the onset *m* on the second copy, replacing an existing onset if there is one (e.g. *dergi* 'journals', *dergi mergi* 'journals or periodicals or magazines'). According to Lewis, this construction cannot be used if the word begins with *m* already (e.g. *müfettişler* 'inspectors'), and a periphrastic construction with *falan* or *filân* is used in its place (*müfettişler falan* 'inspectors and all that lot'). In Abkhaz, suppletive allomorphy comes to the rescue in the comparable situation. Bruening (1997), citing Vaux (1996), describes an Abkhaz echo-word construction which replaces the onset of the second copy with /m/ (*gaʒá-k'* 'fool' → *gaʒák'-maʒák'*); however, if the word already begins with /m/, /č'/ is used instead (*gaʒá-k'* 'secret' → *maʒá-k'-č'aʒá-k'*). This kind of required dissimilation seen in Melodic Overwriting is in some ways reminiscent of the anti-homophony morphological effects described in Section 2.4, which require inputs and outputs, or members of the same paradigm, to differ. Yip (1997, 1998) likens dissimilatory Melodic Overwriting to the kind of conventional poetic rhyme in which identity is required in one prosodic location (e.g. the syllable rhyme) but non-identity is required elsewhere (e.g. the onset of that same syllable); thus *rhyme-time* is a good rhyme but *rhyme-rhyme* is not.

Alderete *et al.* (1999) have argued that some cases of fixed segmentism are, rather than instances of Melodic Overwriting, instead the result of reduction driven by emergent unmarkedness, a phenomenon observed by Steriade 1988 to characterize partial reduplication. In Nupe gerundives (Downing 2004: 90, citing Akinlabi 1997; Smith 1969), an initial CV reduplicant has a fixed [+high] vowel and mid tone, regardless of what is found in the base: *kpi-kpà* 'drizzling', *ji-jákpe* 'stooping', *bi-bé* 'coming'. Insofar as [+high] and mid are the unmarked values for vowels and tone in Nupe, as argued by Akinlabi, the fixed values in the reduplicant can be derived, rather than stipulated. Reduction in partial reduplication is consistent with the hypothesis that partial reduplication typically derives historically from erosion of total reduplication (e.g. Niepokuj 1997). Total reduplication, however, virtually never displays phonological reduction in one copy, as observed in Inkelas (2008a: 379–380).

4.2.3 Morphological Character of Reduplicant It is tempting, based on form, to characterize partial reduplication as affixation and total reduplication as compounding. However, there is little morphological evidence for this distinction. Indeed, some phonologists have recruited the affixation/compounding distinction to account for phonological size differences *within* partial reduplication, terming reduplicants which are syllable-sized or smaller "affixes" and those which are foot-sized "roots" (Generalized Template Theory; for example, McCarthy and Prince 1994, Urbanczyk 1996). This distinction is generally motivated not by

morphological criteria but by the desire to avoid morpheme-specific reference to prosodic templates. Taking a different view, Inkelas and Zoll (2005) observe that reduplication constructions do sometimes impose morphological restrictions that are independent of prosodic ones, arguing against conflating the two types of restriction; they point (in Chapter 2) to the distinction between constructions which specifically double affixes, regardless of size, vs. those that double roots or stems, as evidence that reduplication targets morphologically defined constituents and then imposes phonological shape requirements on the output of doubling.

5 When Phonology is Morphology: Realizational Morphology and Morphologically Conditioned Phonology

Morphologically conditioned phonology overlaps significantly with what has been called "realizational" or "process" morphology, an observation made by, among others, Ford and Singh (1983, 1985), Poser (1984), Dressler (1985), Singh (1996), S. Anderson (1992), Bochner (1992), Orgun (1996), Inkelas (2008b).

To illustrate these phenomena and their overlap, consider two cases of final consonant deletion. The first is a well-known process of subtractive morphology in Tohono O'odham, discussed by S. Anderson (1992), citing Zepeda (1983), and Yu (2000), citing Zepeda (1984). In Tohono O'odham, perfective verbs are derived from imperfectives through the deletion of a final segment (*síkon* 'hoe object' → *síko* (-PERF); *híwa* 'rub against object' → *híw* (-PERF) (Yu 2000: 129–130). This fits the standard description of realizational morphology because there is no other morphological exponent of the perfective.

Now consider the diminutive suffix /-CIk/ in Turkish, discussed in Section 2.1, which triggers an optional process of stem-final velar deletion (/bebek-CIk/ → *bebecik* 'baby-DIM', /köpek-CIk/ → *köpecik* 'dog-DIM') (Lewis 1967: 57) that applies before no other similar suffix. This would standardly be described as a morphologically conditioned phonological rule, because the morphological category of diminutive is marked overtly by the suffix. The operative intuition is that the suffix *-CIk* marks diminutive morphology, while the consonant deletion is just incidental.

This practical distinction between phonology as primary exponent and phonology as secondary concomitant does not always hold up. Sometimes it is difficult, even unproductive, given several exponents of a given morphological construction, to identify which is the primary (morphological) exponent and which are the phonological accompaniments. In Hausa, for example, tone melody replacement can serve as the sole mark of a morphological construction (8a), and so can overt affixation (8b). When both co-occur (8c), is tone melody replacement considered realizational morphology, such that the words in (8c) exhibit two morphological exponents of pluralization, or is tone melody replacement subjugated in (8c) to morphologically conditioned phonology? Page numbers refer to Newman (2000):

(8) a. No affixation; tone replacement (imperative formation)

ká꞉mà꞉	→	kà꞉má꞉	'catch (!)'	(267)
bíncìké꞉	→	bìncìké꞉	'investigate (!)'	(267)
nánné꞉mó꞉	→	nànnè꞉mó꞉	'seek repeatedly (!)' (< *né꞉mó꞉* 'seek')	(263)

b. Overt suffixation, no tone replacement (various)

dáfà꞉	→	dáfà꞉-wá	'cook-VBL.N' -LH	(699)
gàjé꞉ré꞉	→	gàjé꞉r-ìyá꞉	'short-FEM' -LH	(212)
hù꞉lá꞉	→	hù꞉lâ-r	'hat-DEF' -L	(144)

c. Overt affixation *and* tone replacement (various plural classes)

má꞉làm	→	mà꞉làm-ái	'teacher-PL.' -LH	(434)
hù꞉lá꞉	→	hú꞉l-únà꞉	'cap-PL.' -HL	(444)
tàmbáyà꞉	→	támbáy-ó꞉yí꞉	'question-PL.' -H	(432)

One possible way to avoid analytical ambiguity in the case of (8c) would be to reduce all phonological effects other than overt segmental affixation to morphologically conditioned phonology, reanalyzing apparent cases of realizational morphology as zero derivation accompanied by morphologically conditioned phonology. Alternatively, one could try to reduce all morphologically-specific phonological effects to realizational morphology, analyzing the data in (8c) as instances of "extended exponence," the multiple marking of a morphological category (for example, Matthews 1972; Stump 1991). Multiple exponence of overt morphology is a common enough phenomenon; in Hausa, for example, the formation of class 13 noun plurals involves suffixation (of *-e*), (LH) tone replacement, *and* reduplication, for example, *kwánà꞉* 'corner, curve' → *kwàné-kwàné* (PL) (Newman 2000: 458). Harris (2002, 2008b) has argued that circumfixes result diachronically from an earlier stage of multiple affixation, or multiple exponence.

5.1 Theoretical Approaches to Realizational Morphology and Morphologically Conditioned Phonology

The literature on morphologically conditioned phonology, primarily represented by item-based approaches, has had little to say about realizational morphology, despite the obvious formal similarities between the two phenomena. One reason for this is that much of the most influential literature on morphologically conditioned phonology, going back to Kiparsky's (1982b) theory of Lexical Morphology and Phonology (LMP; see also Kiparsky 1984; Mohanan 1986), focuses on phonological patterns common to the morphology of a certain stratum. Both LMP and its successor, Stratal Optimality Theory (Kiparsky 2000, 2008a, to appear), assume a grammatical architecture in which the morphological constructions of each language cluster into a small, possibly universally fixed number of sets ("levels," "strata"), each internally uniform in its phonological patterning, which are totally ordered. In stratal theories like these it is necessary to know only the stratum to which a morphological construction belongs to predict which phonological patterns it will conform to; which stratum a construction belongs to is predictable

from its place in the morphology, that is, whether it is an "early" or "inner" affix as opposed to a "late" or "outer" one.

Because stratal theories focus on commonalities, they are not suited to the description of phonology which is unique to a particular morphological category; very narrowly conditioned phonological effects have to be set aside and treated as exceptions within a stratum, rather than constituting their own individual stratum. For example, in English both *-ible* and *-ive* trigger spirantization on a preceding consonant, an unambiguously stratum 1 effect (*divide, divis-ible, divis-ive*). However, *-ible* triggers voicing while *-ive* does not. This distinction cannot be captured by stratal assignment but must be tied to individual suffixes using exception features or other mechanisms besides strata.

Because, by its nature, realizational morphology is also narrowly tied to individual morphological contexts, stratal ordering theories do not lend themselves to the description of realizational morphology any more than they are suited to capturing idiosyncratic morphophonology.

A middle ground which can capture those morphophonological generalizations sought by stratal theories but also describe highly morphologically-specific phonological patterns is represented by cophonological models. These, as discussed in Section 2, associate each morphological construction (affixation, compounding, zero-derivation) with its own phonological mapping. The cophonological approach eliminates the "too many analyses" problem by using exactly the same mechanism to handle realizational morphology and morphologically conditioned phonology. A phonological alternation specific to a particular affix is included in the cophonology that is unique to that affix. Realizational morphology is accomplished by the cophonology of what might otherwise be described as phonologically null morphological constructions. Whether a construction is "null" or not, that is, whether or not it is associated with an overt affix, is in cophonology theory almost incidental. In this way cophonology theory resembles the approach of Bochner (1992), in which phonological patterns are part and parcel of the description of the rules relating words in a paradigm. Cophonology theory is not limited to enumerating idiosyncracies; as Anttila (1997a, 2002a) has demonstrated, its inheritance architecture also gives it the ability to posit meta-constructions like "word," "stem," or "stratum," with associated cophonological restrictions inherited by the member constructions, to capture generalizations holding across all constituents of a certain type.

6 When Phonology and Morphology Diverge: Nonparallelism Between Phonological and Morphological Structure

The domains of word-internal phonological patterns are generally coextensive with the morphological sub-constituents of a word; for this reason, phonology provides strong evidence about the morphological structure of a word. However,

there can be mismatches, that is, situations in which phonological domains are not matched with morphological sub-constituents. In some cases the phonological domain – prosodic root, or stem, or word – is a sub-portion of a word (see e.g. Booij 1984; Sproat 1986; Inkelas 1990; Booij and Lieber 1993, among many others). Three situations stand out in this regard: compounding, the distinction between cohering and non-cohering affixes, and reduplication of an internal prosodic stem.

The literature on the phonology of compounding constructions has often drawn attention to a distinction between compounds that behave phonologically like one word and those that behave phonologically like two words. In the 1980s this difference was attributed to prosodic structure which is loosely related to but exists independently of morphological and syntactic structure. Nespor and Vogel (1986) proposed that while Greek compounds form a single prosodic word and thereby receive one stress, for example, *kúkla* 'doll' + *spíti* 'house' → [*kuklóspito*] 'doll's house' (p. 112), the members of Hungarian compounds form separate prosodic words and retain their own lexical stresses: [könyv] [tár] 'book collection' (p. 123). In Malayalam, simple sub(ordinate) compounds, with head-modifier semantics, form a single domain for accentuation, whereas the members of simple co(ordinate) compounds, with coordination semantics, form separate domains for accentuation. Sproat (1986) and Inkelas (1990) proposed that this difference in behavior could be attributed to different prosodic structure, though Mohanan (1995) later countered this argument with evidence from complex compounds with three or more members. In a detailed study of Indonesian, Cohn (1989) documents a stress difference between two constructions that concatenate stems. Head-modifier compounds impose stress reduction on one member (*polúsi* 'pollution' + *udára* 'air' = *polùsi udára* 'air pollution', p. 188), suggesting that they are competing for prominence within a single phonological word, while total reduplication constructions maintain two equal stresses (*minúman* 'drink', *minúman-minúman* 'drinks', p. 184). Cohn attributes the latter pattern to the fact that total reduplication consists of two prosodic words. Itô and Mester (1996a) point to a similar distinction in Japanese, in which stem-stem compounds form one prosodic domain, word-word compounds form two domains, and stem-word compounds "type-shift," by means of a principle of Prosodic Homogeneity (p. 38), to pattern like word-word compounds (see also Han 1994 on Korean).

Perhaps even more interesting than prosodic differences across types of compounds are comparable differences in affixed words. Booij (1984) was one of the first to highlight the distinction between "cohering" and "non-cohering" affixes and to model the distinction using prosodic structure: cohering affixes form a single prosodic word with the base of affixation, while non-cohering affixes form a separate prosodic domain. In Dutch, for example, suffixes are either cohering, meaning they syllabify with and join into a prosodic domain with the stem they combine with, or non-cohering, meaning they create a separate prosodic domain. Nonnative suffixes in Dutch are all of the non-cohering type. The difference between the two types of suffix is illustrated with this minimal pair of suffixes both of which are equivalent to English "-ish": *rood-achtig* [roːt.ɑx.təx] and *rod-ig*

[roɪ.dəx] 'reddish' (Booij 1984 :152). As Booij (1984, 2001, 2002: 172) observes, non-cohering Dutch suffixes like *-achtig* allow deletion under identity: if two words ending in *-achtig* are coordinated, the suffix is omissible from the first conjunct: *storm-achtig en regen-achtig* 'stormy and rainy' can also be realized as *storm en regen-achtig* (1984: 151). Deletion under identity is also found in compounding constructions: *wespen-steken en bije-steken ~ wespen en bije-steken* 'wasp and bee stings' (1984: 146). Non-cohering affixes and members of compounds form individual prosodic words, explaining their parallel behavior. By contrast, cohering suffixes like *-ig* cannot be omitted under identity: *blau-ig en rod-ig* 'blueish and reddish', but **blau en rod-ig* (1984: 149). On deletion under morphological identity in other languages, see, for example, Vigário and Frota (2002), Orgun (1996).

Evidence for the accessibility to "later" processes of word-internal prosodic stems is found in reduplication. In a number of cases, a late morphological process of reduplication targets the root, even if the root has already undergone significant affixation. Aronoff (1988) refers to these as "head operations," and Booij and Lieber (1993) propose that they involve reference to a prosodic stem, which corresponds closely if not exactly to the morphological root. Inkelas and Zoll (2005) cite the example of Chumash, which has what Applegate (1972: 383–384) characterizes as a very late process of reduplication, conferring the meaning of a repetitive, distributive, intensive, or continuative force. Chumash reduplication targets a sub-constituent of the word which Inkelas and Zoll term the prosodic stem. The prosodic stem always contains the root, along with any preceding prefixes of the type Applegate (1972) identifies as reduplicating, and which Inkelas and Zoll (2005) analyze, in Booij's terms, as cohering. For example, the root-adjacent prefix in *k-sili-*{*pil-*<u>*wayan*</u>} 'I want to swing', in which curly brackets demarcate the prosodic stem and the root is underlined, is cohering and participates in reduplication: *ksili*{*piw-piwayan*} (Applegate 1972: 387). By contrast, the prefixes in *s-am-ti-*{<u>*lok'in*</u>} 'they cut it off' are non-cohering and do not reduplicate: *samti*{*lok-lok'in*} (Applegate 1972: 387). Evidence that what reduplicates is a prosodic stem, occupied by the root and joined by cohering prefixes, is that the prosodic stem is subject to a typical stem-shape constraint; it must be consonant-initial. Onset consonants are not required of Chumash roots or prefixes, many of which are vowel-initial. But prosodic stems must be consonant-initial. As a result, even an otherwise non-cohering prefix will contribute its final consonant to a following prosodic stem, as shown by reduplicated forms such as *s-iy-ak*{*t-aqu-*<u>*smon*</u>} → *siyak*{*taq-taqusmon*} 'they come to gather it' (Applegate 1972: 388). Parallel phenomena, documented in Inkelas and Zoll (2005), occur in Tagalog (see also Booij and Lieber 1993) and Eastern Kadazan (Hurlbut 1988). An alternative analysis of the Chumash and Tagalog phenomena is offered within Base-Reduplicant Correspondence Theory (BRCT) by McCarthy and Prince (1995), who propose that the reduplicant is not infixing but is instead prefixed directly to the material that is copied. On their account, the copying of the final consonant of a prefix *preceding* the reduplicant is the result of morphological fusion between the prefix consonant and the VC reduplicant and "back-copying" of the result to the base of reduplication: *s-iš-*$_{RED}$*-expeč* → *s-i-šex*$_{RED}$*-šexpeč*$_{BASE}$ (with backcopying of the *iš* -final *š* to the base).

Inkelas and Zoll (2005) argue against this account of Chumash, in particular, on language-internal morphological grounds. McCarthy and Prince (1995) have, however, identified other apparent cases of backcopying in other languages, and backcopying in general remains a viable analysis within BRCT.

Returning to mismatches between morphological constituent structure and prosodic structure, there is also strong evidence that word-sized prosodic domains can include material outside of the morphological or lexical word, clitics being the most obvious example. It has been widely argued that clitics are phonologically defective syntactic terminal elements, having to join with another (non-clitic) syntactic terminal element to form a single prosodic word (e.g. Inkelas 1990; Halpern 1992; Booij 1996). A question of considerable current interest is whether prosodic word structure can be recursive; see Peperkamp (1996), Itô and Mester (2003a), and Kabak and Reviathidou (2009), among others.

7 Summary

The phonology-morphology interaction sheds light on word-internal structure and on the ability for relatively unnatural phonological alternations to be productive, at least within a given morphological niche. Both realizational morphology and morphologically conditioned phonology operate in the same domains and manipulate the same structural elements. The many related phenomena constituting the phonology-morphology interface are central to word formation in virtually all languages, and must therefore be taken seriously by morphologists and phonologists, especially those seeking to reduce synchronic morphological patterns to syntax, or synchronic phonological patterns to universal phonetic motivations.

NOTES

1 This is the approximate generalization, as stated by Kenstowicz; the actual picture is more detailed, in ways not material to the point made here. See, for example, Butt and Benjamin (2008).

2 Vaux (p. 252) analyzes the definite suffix as underlyingly /-n/ and attributes the schwa allomorph to rules of epenthesis and consonant deletion; however, as the n~ə alternation is specific to the definite, and most researchers would probably classify this as suppletive allomorphy.

4 Quantity

STUART DAVIS

1 Introduction

Quantity plays an important role in the analysis of a wide variety of phonological and morphological phenomena in many languages including the analysis of word stress, tone, compensatory lengthening, shortening processes, minimal word requirements, templatic restrictions, and allomorphy selection. These phenomena frequently distinguish between syllables that are short (or light) from those that are long (or heavy). In the first section of this chapter we introduce the theory of moraic phonology of Hayes (1989a), a representational theory of quantity. In subsequent sections we overview a number of issues that emerge from moraic theory. In our discussion of these issues we will refer to the wide variety of processes in which quantity plays a role.

In modern studies of phonology, the term "quantity" refers to either segmental duration or syllable weight. With respect to segmental duration, quantity differences among segments are said to be phonemic in languages that contrast a long and short form of a vowel of the same quality and in languages that contrast a geminate versus non-geminate consonant. The Japanese examples in (1) illustrate both types of contrasts.

(1) Japanese quantity contrasts (Tsujimura 2007)[1]

a.	[su]	'vinegar'
b.	[su:]	'inhale'
c.	[saka]	'hill'
d.	[sak:a]	'author'

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

Phonetically, though, a single phoneme may be pronounced longer or shorter, depending on the nature of the environment in which the segment occurs. For example, cross-linguistically a vowel tends to have longer duration before a voiced consonant than before a voiceless one (Chen 1970) and tends to be longer if it precedes a fricative as opposed to a stop (Peterson and Lehiste 1960; House 1961). However, such phonetic environmental differences are not relevant for processes like stress placement that often distinguish syllables with long vowels. Thus, for instance, we do not find rules of stress assignment that place stress on a syllable containing a vowel preceding a voiced consonant or a fricative.[2] On the other hand, phonemic contrasts in quantity as in (1) often play an important role in the phonology of languages that have such contrasts.

With the advent of generative phonology, the major issue concerning quantity has been the nature of its phonological representation.[3] Chomsky and Halle (1968) use the feature [±long] to characterize segmental quantity. Under such a characterization, the difference between the Japanese words in (1a) and (1b) is that [u] would have the feature [+long] in the latter and [–long] in the former. However, it was subsequently noted (e.g. Leben 1980) that this representation was insufficient, since long consonants can behave like a sequence of two segments for certain phenomena. Further, inalterability effects were noted by such researchers as Kenstowicz and Pyle (1973), Schein and Steriade (1986) and Hayes (1986) whereby long segments seemed to be immune to certain phonological processes that shorter segments of the same quality underwent. Such observations motivated an autosegmental representation of segmental quantity, in which long vowels and geminates are linked to two slots on a timing or prosodic tier while a short vowel or singleton consonant is linked to one slot. There has been much literature on the nature of this timing or prosodic tier (see Kenstowicz 1994a; Broselow 1995; and Hermans 2006 for overviews). A common view of this tier was that it either consisted of CV-slots (e.g. McCarthy 1979a, 1981; Halle and Vergnaud 1980; Clements and Keyser 1982) or X-slots (Levin 1985). This is shown in (2) and (3) below where (2) gives both the CV-tier and X-tier representation of the Japanese word in (1b), and (3) gives the CV-tier and X-tier representation of the Japanese word in (1d).

(2) a. CV-tier representation of [su:]

b. X-tier representation of [su:]

(3) a. CV-tier representation of [sak:a]

b. X-tier representation of [sak:a]

Notice that in (2) and (3) the prosodic shape of the word is encoded segmentally as a sequence of CV-slots or X-slots. Because of this encoding, the CV or X-tier is often referred to as a prosodic tier.

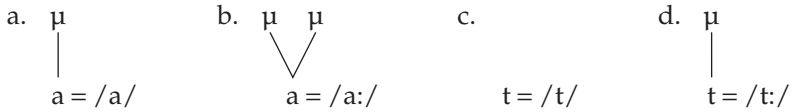
In a highly influential paper, Hayes (1989a) rejected the segmental nature of the prosodic tier and argued instead for its characterization as moraic. Hayes's main argument against a segmental CV or X-tier is that it fails to properly identify which types of segmental deletions lead to compensatory lengthening. To see what is at issue, consider Turkish compensatory lengthening, discussed by Sezer (1986) and more recently by Hermans (2006). In Turkish, the phoneme /v/ can optionally delete in certain postvocalic environments as shown in (4).

(4) Turkish optional /v/ deletion (Sezer 1986: 228)

a. [davul] – [daul] 'drum'
b. [savmak] – [sa:mak] 'to get rid of'

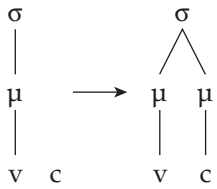
In a CV-tier or X-tier representation, there is no non-stipulative explanation for why the deletion of the consonantal phoneme /v/ should lead to compensatory lengthening in (4b) but not in (4a). Under Hayes's (1989a) theory the difference between examples like those in (4a) and (4b) is that in (4b) the deleted /v/ is a moraic segment in that it is in the coda of the (first) syllable, but in (4a) it is not moraic since it is in the onset of the (second) syllable. Hayes (1989a) makes a strong case that compensatory lengthening involves the loss of a moraic segment without the deletion of the mora. While the notion of mora in current phonological theory goes back at least to Trubetzkoy's (1939) discussion of syllable quantity,[4] Hayes (1989a) develops a formal theory of moraic phonology in which the prosodic tier is characterized as moraic.[5] Specifically, in Hayes's theory, a short vowel is underlyingly monomoraic while a long vowel is bimoraic. With respect to geminate consonants, a geminate consonant differs from a short consonant in that the geminate is underlyingly moraic while a short consonant is non-moraic. Sample moraic representations are given in (5) where (5a) shows a short vowel, (5b) a long vowel, (5c) a short consonant, and (5d) a geminate.

(5) Underlying moraic representation (Hayes 1989a)



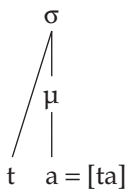
In Hayes's theory a (non-geminate) coda consonant is not underlyingly moraic. Rather, in some languages a coda consonant acquires moraic status by the rule of Weight-by-Position shown in (6). (We indicate the syllable with the symbol σ.)[6]

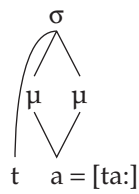
(6) Weight-by-Position

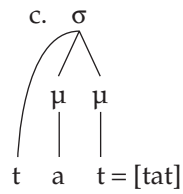


In (7) we provide examples of syllabification with the surface moraic structure shown.

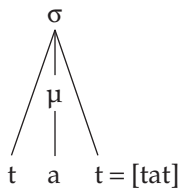
(7) Surface syllabification with moraic structure[7]

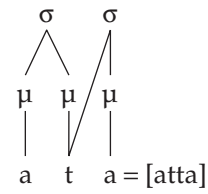
a. 

b. 

c. 

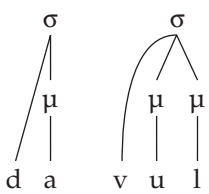
or

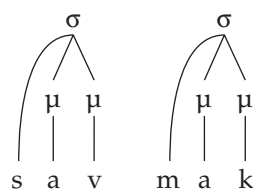
d. 

e. 

While the sequence of phonemes in (7c) and (7d) is identical, they differ in that the rule of Weight-by-Position in (6) has applied to (7c) but not to (7d). That is, the coda consonant in (7c) is moraic while the coda in (7d) is not. (7e) shows the mora structure and syllabification of a form with a geminate consonant, which on Hayes's theory is underlyingly moraic.[8] Now, given this view of moraic structure, it becomes quite clear why Turkish /v/ deletion in (4b) triggers compensatory lengthening while that in (4a) does not. In (8) we show the syllabification and moraic structure of the two words in (4) before compensatory lengthening takes place.

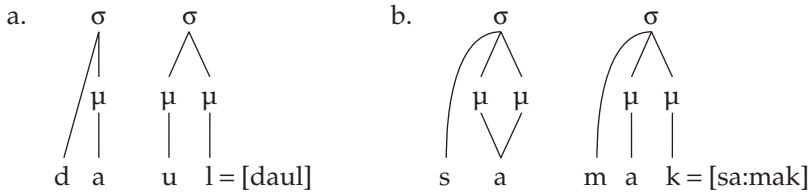
(8) Mora and syllable structure of (4)

a. 

b. 

In Hayes's theory, the deletion of /v/ in (8b) as reflected in (4b) results in compensatory lengthening while the deletion of /v/ in (8a) as reflected in (4a) does not trigger it. Assuming that Weight-by-Position applies in Turkish, the deletion of a moraic segment triggers compensatory lengthening in order to preserve the mora, while deletion of a non-moraic segment does not result in compensatory lengthening, since no mora is involved. (9a) and (9b) show the two Turkish words after /v/ deletion has applied, with compensatory lengthening in (9b).

(9) Compensatory lengthening as mora preservation



As seen by examining (8) and (9), compensatory lengthening preserves mora structure despite the deletion of the segment. Consequently, Hayes's theory makes a prediction that it is only the loss of a moraic segment that can lead to compensatory lengthening. To the extent that this is correct, it provides a strong argument for the moraic representation of segmental quantity and for the moraic nature of the prosodic tier.[9]

The moraic representation of segmental quantity in which certain segments are either underlyingly moraic or surface as moraic (through Weight-by-Position) plays an important role in the characterization of syllable quantity (i.e. syllable weight) in current phonological theory. This is because it provides a formal distinction between a light syllable and a heavy syllable. A light syllable is one that is monomoraic whereas a heavy syllable is bimoraic (or greater). Thus, the syllables in (7a) and (7d) are light or monomoraic whereas those in (7b) and (7c) are heavy or bimoraic. (The issue of geminate consonants as in (7e) is ignored here but will be discussed in detail in Section 5.) The difference between light and heavy syllables is of significance for a wide variety of phonological and morphological processes including stress assignment (e.g. languages in which stress picks out heavy syllables), tonal realization (e.g. languages in which contour tones can be realized only on heavy syllables), compensatory lengthening, closed syllable shortening, minimal word effects (e.g. languages that require a bimoraic word minimum), and morphological phenomena such as templatic morphology (McCarthy and Prince 1986; Crowhurst 1991; Broselow 1995), allomorphy selection, and mora augmentation (Davis and Ueda 2006, Fitzgerald 2012).

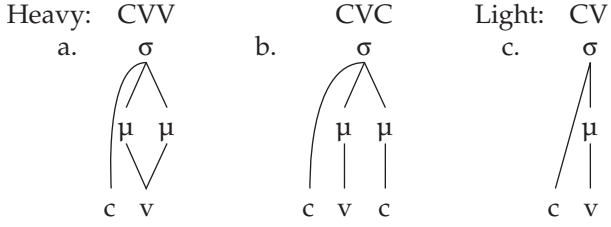
Given the moraic structure shown in (7a–d), we expect to find at least two types of systems with respect to syllable weight in languages in which syllable weight plays a role. These are shown in (10).[10]

(10) Syllable weight systems:

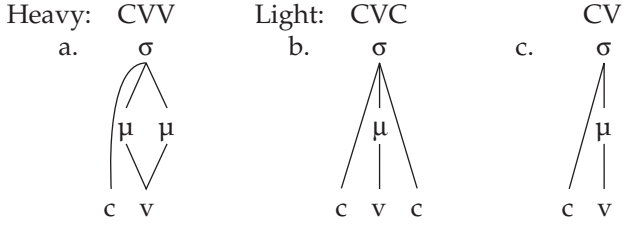
a. *Heavy*	*Light*	b. *Heavy*	*Light*
CVV	CV	CVV	CV
CVC			CVC

The system in (10a) reflects languages where Weight-by-Position applies to coda consonants, so that both CVV and CVC syllables are treated as heavy. The system in (10b) reflects languages in which Weight-by-Position does not apply, so that only CVV syllables are treated as heavy while CVC syllables pattern with light CV syllables.[11] The system in (10a) is shown with its moraic structure in (11), and the system in (10b) is shown with its moraic structure in (12). (We use "c" and "v" below to indicate a consonant phoneme and a vowel phoneme, respectively.)

(11) Moraic structure for Heavy CVV, CVC vs. Light CV



(12) Moraic structure for Heavy CVV vs. Light CVC, CV



The system of moraic structure as presented in (11) and (12), reflecting the two types of weight systems illustrated in (10), is a commonly accepted view of moraic representation and syllable weight. Much current work on syllable quantity takes (10)–(12) as a starting point even if to argue against moraic representations or specific aspects of moraic theory. In the remainder of this chapter, we will take (10)–(12) as a basis for an overview of various issues and elaborations that have been brought up in the literature on the moraic theory of quantity since Hayes's seminal article on the topic, such as conditions on coda weight, the possibility of more than two degrees of syllable weight, the apparent inertness of syllable onsets to syllable weight, the underlying moraicity of geminate consonants, and weight inconsistency whereby, for example, within one language CVC syllables can act as both heavy and light. The conclusion will include a brief discussion of the phonetics of syllable weight.

2 Coda Weight

In the moraic theory that is presented in (10)–(12), no distinction is made between coda consonants of different types: either codas are moraic (11b) or not (12b). However, work by Zec (1988, 1995b) that harks back to Trubetzkoy (1939): 170–171; pagination as in Baltaxe (1969) posits that languages can impose a minimal sonority threshold on what can be a moraic segment. In support of this, Zec presents evidence from a variety of languages in which a CVC syllable closed by a sonorant consonant acts as heavy while a CVC syllable closed by an obstruent acts as light. In these languages there is a minimal sonority threshold for moraicity: only vowels and sonorant consonants can be moraic. Consequently, Zec argues that in addition to the weight systems given in (10) some languages have the weight system in (13).[12] As she notes, this division was also discussed by Prince (1983).

(13) Syllable weight distinction based on sonority (S = sonorant consonant and O = obstruent)

Heavy	*Light*
CVV	CV
CVS	CVO

Zec cites Lithuanian as an example of a language that has the syllable weight system in (13). She gives two pieces of evidence in support of this division. In (14), we show Zec's (1995b) examples of a morphological ablaut process in Lithuanian that has the effect of lengthening the vowel of the root in certain forms of the verb. (The double vowel representation of length in (14) follows Zec.)

(14) Verbal ablaut in Lithuanian (Zec 1995b: 100–102)

	Verb stem	*Infinitive*	*Gloss*
a.	tup	tuup-ti	'perch'
b.	dreb	dreeb-ti	'splash'
c.	vag	voog-ti	'steal'
d.	vir	vir-ti (*viir-ti)	'boil'
e.	mir	mir-ti (*miir-ti)	'die'
f.	kar	kar-ti (*kaar-ti)	'hang'
g.	kau	kau-ti	'beat'

Ablaut applies to the infinitive forms in (14a–c) lengthening the root vowel. Vowel lengthening does not occur in (14d–g). The verb stems in (14) are similar to one another in that they all consist of a single closed syllable (assuming that the off-glide in (14g) also closes the syllable and functions as a coda).[13] The key difference between (14a–c) and (14d–g) is that the stem syllable ends in a sonorant in the latter but in an obstruent in the former. If a sonorant consonant in the stem syllable is viewed as being moraic then the lack of vowel lengthening in (14d–g)

can be understood as the avoidance of trimoraic syllables. In (14a–c) the stem-final obstruent is not moraic so the vowel can lengthen without creating a trimoraic syllable. Independent support for this moraic difference between coda sonorants and obstruents comes from a restriction on the occurrence of the circumflex accent in Lithuanian. The circumflex accent is an accent type in Lithuanian that is associated with a rising tone. Zec (1988, 1995b) observes that this accent type only occurs on syllables with a long vowel or closed by a sonorant consonant. It does not occur on a syllable closed by an obstruent. This suggests that the circumflex accent can only be realized on a bimoraic syllable, and as Zec (1995b: 98) proposes, Lithuanian has a minimal sonority threshold limiting moraic elements to those that are [+sonorant]; a coda obstruent does not count as moraic.

Zec's discussion of Lithuanian seems to support the weight distinction in (13) between sonorant and obstruent codas, but several issues remain unresolved. One important question that Zec (1995b) raises is whether other types of restrictions on coda consonants, referred to as edge constraints by Zec, are reducible to sonority threshold restrictions on moraicity. Zec argues convincingly that these two types of constraints affecting coda consonants are different. This is made clear by her discussion of the Australian language Lardil (Hale 1973). In Lardil, codas are restricted so that only (non-dental) coronal consonants can close a syllable, but the coda does not count as moraic since CVC words do not satisfy a bimoraic minimality requirement. Thus, according to Zec, Lardil exemplifies a language that has edge restrictions on coda consonants, but they are not reducible to restrictions on moraicity. Zec (1995b: 111–112) contrasts Lardil with the Tanoan language Kiowa which limits the coda to the consonants /p t m n l/. This set includes the sonorant consonants plus two obstruents. There are other obstruents in the Kiowa inventory (such as fricatives and dorsal stops) but only /p t/ can occur in coda position among the obstruents. Zec notes that in Kiowa, long vowels shorten in closed syllables regardless of whether the syllable is closed by a sonorant consonant or an obstruent. Assuming that closed syllable shortening occurs to prevent trimoraic syllables, Zec maintains that all codas in Kiowa are moraic, which includes both sonorant and obstruent consonants. Zec thus concludes that in Kiowa all codas can be moraic (i.e. there is no sonority threshold on moraic segments), but there are independent edge constraints that prohibit dorsal consonants and fricatives in the coda. If we bring together Zec's discussion of Lithuanian with that of Lardil and Kiowa, we see three different examples of the relation between coda restrictions and coda moraicity. In Lithuanian, there are no major restrictions on what can appear as a single coda; however, only sonorant codas are moraic. In Lardil, edge constraints limit the nature of the coda to (non-dental) coronals, but only vowels can be moraic; that is, the sonority threshold prevents consonantal elements from being moraic. In Kiowa, edge constraints allow only /p t m n l/ as codas, but there are no sonority threshold constraints on moras; any consonant (sonorant or obstruent) can constitute a moraic coda. Finally, languages may witness a fourth possibility where there are few or no edge constraints restricting the nature of codas and where there is no sonority threshold on what can be moraic. Latin can be cited as an example since there are no (relevant) edge constraints on what can be a coda and any type of consonant in the coda is moraic.

There is another issue that emerges from Zec's (1988, 1995b) proposal. Zec (1995b) defines sonority based on the major class features, basically [±consonantal] and [±sonorant]. This gives a division between obstruents, sonorants, and vowels. However, there seem to be languages where, in addition to sonorants, high sonority obstruents are also included in the class of moraic segments. One example is Tiv which is mentioned by Zec (1995b), though not discussed, as having the weight division in (13). However, an examination of Tiv (see Pulleyblank 1988, in particular) reveals that the consonants that can appear in coda position are sonorants and voiced fricatives. While it may be possible to interpret voiced fricatives as sonorant, one can just as well note that since voiced fricatives are the most sonorant of the obstruent types in Tiv, the sonority threshold on what can be moraic in Tiv would just be any phoneme with the sonority value of at least a voiced fricative. While such cases are not discussed by Zec, they are not incompatible with her theory, assuming a finer division of the sonority hierarchy than that allowed by the major class features [±consonantal] and [±sonorant].[14] What would not be expected under Zec's theory are languages in which both obstruents and sonorants appear in coda position but where only syllables closed by obstruents make a syllable heavy.

A related issue that can be raised regarding Zec's proposal is whether there are restrictions on coda moraicity that are not sonority related, a problem highlighted by comparing Japanese and the Australian language Ngalakgan. Japanese obeys the Coda Condition in the sense of Itô (1986), whereby coda consonants must be assimilated or place-linked to the following onset, but where coda consonants are clearly moraic (e.g. Kubozono 1999). It is possible to interpret this as a combination of an edge constraint (namely the Coda Condition) with a sonority threshold that allows any segment type (obstruents included) to be moraic. To put it another way, Japanese would have the weight division in (10a) in which any coda consonant can be moraic along with the independent Coda Condition that bans coda consonants from having their own place features.

In light of this, it is interesting to consider the Australian language Ngalakgan discussed by Baker (1997, 2008) and Davis (2003). Consider the pattern of primary stress reflected by the data in (15) taken from Baker.

(15) Ngalagkan (Capitalized consonants indicate retroflex sounds.)

a.	cíwi	'liver'
b.	céraTa	'women's ceremony'
c.	páRamùnu	'sand goanna'
d.	cálapir	'red ant (species)'
e.	kúpuy	'sweat (n.)'
f.	purúTci	'water python'
g.	kipíTkuluc	'frogmouth (bird)'
h.	miRárppuʔ	'crab'
i.	puTólkoʔ	'brolga (bird)'
j.	màcapúrka	'plant (sp.)'
k.	LáRkurca	'vine (sp.)'
l.	cálpurkic	'fish (sp.)'

m.	kaNTálppuru	'female plains kangaroo'
n.	cákanta	'macropod (sp.)'
o.	ŋúruNTuc	'emu'
p.	ŋóloŋkoʔ	'eucalyptus'
q.	ŋámuccùlo	'subsection term'
r.	cápatta	'tortoise (sp.)'
s.	móLoppoL	'catfish (sp.)'
t.	ŋáNaʔpay	'and moreover'

The data in (15a–l) show that primary stress in Ngalakgan falls on the leftmost (non-final) heavy syllable; otherwise, it falls on the initial syllable. One way of analyzing these data is to posit that a coda consonant is moraic in making a syllable heavy be it an obstruent (15f–g) or a sonorant (as in 15h–l). The data in the second column (15m–t) show that the leftmost closed syllable (underlined) fails to attract primary stress. However, the coda in the leftmost closed syllable in (15m–t) is one of three types: in (15m–p) the coda is a nasal homorganic to the following onset; in (15q–s) it is the first part of a geminate consonant, and in (15t) the coda is a glottal stop. What these three coda consonant types have in common is that they do not have their own place features: the place features are either shared with the following onset, or, in the case of the glottal stop in (15t), there is a lack of place features altogether. If we assume that the closed syllables in (15f–l) attract stress because they are bimoraic, then it would seem that Ngalakgan has a restriction that requires moraic elements to have independent place features (i.e. not shared with a following onset). Thus, while a wide variety of segment types can appear in the coda in Ngalakgan, only those having independent place features surface as moraic (but see Baker (2008) for a different interpretation). The Ngalakgan case as analyzed here would be problematic for Zec's theory since it entails a restriction on moraicity that is not sonority related. (See Ní Chiosáin (1990) for discussion of western Irish, wherein codas with shared place features do not act as moraic.) This suggests that the relationship between coda restrictions and coda moraicity is still in need of further study.[15]

3 Multiple degrees of weight

In the discussion presented so far, there has been a tacit assumption that there are only two degrees of syllable quantity: light/monomoraic vs. heavy/bimoraic, though languages may differ along the lines of (10) and (13) as to what is considered light and heavy. We have also mentioned the lack of lengthening in Lithuanian CVS syllables (14d–g) and closed syllable shortening in Kiowa whereby potentially trimoraic syllables consisting of a long vowel and moraic coda are avoided. The avoidance of potentially trimoraic syllables is a rather common phenomenon cross-linguistically. Consider the example in (16) from two different dialects of Arabic where a variety of research has shown that (non-final) coda consonants pattern as moraic (e.g. Kiparsky 2003b).

(16) Avoidance of trimoraic syllables in Arabic dialects (see Broselow 1992 for an overview).

a.	Cairene Arabic:	/baab + na/	–	[báb.na]	'our gate'
b.	Meccan Arabic:	/baab + na/	–	[báa.ba.na]	'our gate'

As seen in (16) both Cairene and Meccan Arabic avoid the potentially trimoraic parse of the first syllable of /baab +na/ as [báab.na]. The dialects, though, differ in how they avoid the trimoraic parse. While Cairene Arabic favors closed syllable shortening, Meccan Arabic preserves the underlying vowel length by having vowel epenthesis apply between the two consonants to create open syllables, thus avoiding any trimoraic syllable. Examples like Arabic, Lithuanian, and Kiowa suggest that languages tend to avoid having more than the degrees of syllable weight. This issue was discussed at length by Trubetzkoy (1939) who maintained that almost all cases where a language has been claimed to have more than two degrees of distinctive quantity for the syllable involve effects that are phonetic rather than phonological.

Despite a tendency in many languages to avoid trimoraic syllables, there are some languages that seem to allow them. One example discussed by Hayes (1995a: 276) is the Hindi dialect described by Kelkar (1968). In this dialect, primary stress falls on the rightmost heaviest (non-final) syllable of the word. The syllable types are listed in (17).

(17) Syllable types in Hindi (Kelkar 1968; Hayes 1995a)

a.	superheavy:	CV:C, CVCC
b.	heavy:	CV:, CVC
c.	light:	CV

The syllable types in (17) can be interpreted as reflecting different quantities with superheavy syllables being trimoraic, heavy syllables bimoraic, and light syllables monomoraic. Stress falls on the rightmost (non-final) syllable with the most moras. In a constraint-based framework, one can view the difference between a language like Hindi allowing trimoraic syllables and Cairene Arabic, which disallows such syllables, as the applicability of a constraint against trimoraic syllables, as in Broselow (1992). This constraint would be inviolable in Cairene Arabic but not in the Hindi dialect described by Kelkar (1968). Thus, there do seem to be cases of more than two degrees of syllable weight.

In addition to cases like Hindi where there are trimoraic CV:C and CVCC syllables, one also comes across cases cited in the literature of what look to be multiple degrees of syllable weight based on vowel quality. Some of these systems have been discussed in detail in the stress literature in Optimality Theory by researchers such as Kenstowicz (1994b), Zec (2003), and De Lacy (2002). One example of such a system discussed by both Kenstowicz and Zec is the Finno-Ugric language Mordwin, in which stress is sensitive to the height quality of a vowel. Essentially, following the interpretation of the Mordwin data in Zec (2003),

primary stress in Mordwin words falls on the leftmost syllable containing a non-high vowel; if the word has only high vowels, then stress falls on the leftmost (initial) syllable. Thus, stress is attracted to a syllable with a non-high vowel. While one may be tempted to analyze the difference between non-high and high vowels in terms of quantity by referencing moraic structure, for example by assigning two moras to non-high vowels and one mora to high vowels, this difference does not seem to be related to quantity. Thus, these researchers do not analyze the Mordwin system (and other similar systems where stress is sensitive to vowel quality) in terms of quantity distinctions. Rather, they are analyzed in terms of preferred quality distinctions in positions of prominence. Kenstowicz (1994b) makes reference in his analysis of Mordwin to constraints that prefer lower vowels as syllable peaks while Zec (2003) references a constraint on the sonority threshold for the head of a foot that in Mordwin must be [−consonantal], [−high].

One of the most complex stress patterns reported in the literature is that found in Nanti, a Kampa language of Peru discussed by Crowhurst and Michael (2005). In this language both vowel quality and quantity are essential in determining the location of primary stress. Nanti has a syllable weight hierarchy somewhat similar to that of Hindi in (17), but vowel quality can play a role in determining stress among syllables of equal weight. Consider first the Nanti stress data in (18) from Crowhurst and Michael (2005) which illustrate the effect of syllable quantity on stress. In examining the data in (18), keep in mind that when there are no overriding factors, stress in Nanti is iambic and iteratively assigned from the left edge of the word, with final syllables normally being extrametrical as in (18a–b). (We do not discuss the issue as to which of the stressed syllables in a word is assigned primary stress. We refer the reader to Crowhurst and Michael for details regarding this and other aspects of the stress system not discussed here.)

(18) Nanti stress (Crowhurst and Michael 2005)

a.	o.kò.wo.gó.te.ro	‘she harvests it’
b.	i.rì.pi.rí.ni.te	‘he will sit’
c.	o.tá.sòŋ.ka.kse.ro	‘she blew on it’
d.	òŋ.ko.wo.gó.te.ro	‘she will harvest it’
e.	pìŋ.kse.ma.wáa.kse.ro	‘you will have listened attentively to it’
f.	jo.bìi.kái.ga.kse	‘they MASC drank’
g.	nóo.ga.ksem.pa.ro	‘I will have consumed it’
h.	i.kà.man.tái.ga.kse.na	‘they MASC told me’
i.	o.sà.ráan.tai.ga.kse	‘they MASC tore it with a purpose’

The data in (18a–b) reflect the normal pattern of iterative iambic footing assigned from the beginning of the word when there are no overriding factors. (18c–e) show that closed syllables (CVC) attract the stress away from a light CV syllable disrupting the iambic pattern. (The only coda consonant allowed in Nanti is a nasal homorganic to the following onset consonant.) This suggests that CVC syllables are heavier than CV syllables. The data in (18f–g) illustrate the effect of

a CVV syllable (i.e. a syllable with a long vowel or diphthong). Such syllables attract stress away from CV syllables. Further, the data item in (18h) is consistent with CVV syllables being heavier than CVC syllables in that stress falls on the fourth syllable (CVV) rather than on the third syllable (CVC). Finally, the comparison of the third and fourth syllables in (18i) demonstrates that CVVC syllables are heavier than CVV syllables. Taken together, the data in (18) suggest the following "weight" hierarchy in (19) for determining which syllable in the Nanti foot receives a stress.

(19) "Weight" (or strength) hierarchy for determining the stressed syllable in the Nanti foot

CVVC > CVV > CVC > CV

This is an interesting hierarchy because both CVV and CVC would be bimoraic under a conventional view of moraic phonology, but nonetheless, CVV patterns as heavier than CVC. Because of this, Crowhurst and Michael take the position that only syllables with long vowels are bimoraic in Nanti and that the coda is not moraic. To account for the apparent heaviness of CVC syllables, Crowhurst and Michael (2005: 57) posit an independent coda strength scale that makes closed syllables (CVC) stronger than CV syllables. Stress is sensitive to this scale as well as to the moraic makeup of the syllable. Consequently, CVV still would be "stronger" than CVC in Nanti since it is bimoraic whereas CVC is monomoraic; a bimoraic syllable is stronger than a monomoraic one.[16]

In addition, vowel quality plays a role in stress assignment when the syllables in a foot have equal strength. Essentially, when syllables of equal strength are in the same foot, stress falls on the syllable having the lower vowel. This is shown in (20).

(20) Vowel quality effect on Nanti stress (Crowhurst and Michael 2005)

a.	à.wo.te.hái.gʒi.ri	'we approached him/them'
b.	nò.gʒi.wo.tá.kse.ro	'I placed it (vessel) mouth down'
c.	noŋ.kàn.tái.ga.kse	'we will have said'
d.	iŋ.ksèn.tá.kse.ro	'he will have pierced it (with an arrow)'
e.	noo.gái.ga.ro	'we ate it'
f.	i.rò.bii.kái.ga.kse	'they MASC will have drunk'

In (20a–b), the first two syllables are light CV; nonetheless, a stress falls on the first syllable rather than the second because in each case the vowel of the first syllable is lower than the vowel of the second syllable. In (20c–d) the first two syllables are CVC; here, in each of these words, a stress is on the second syllable, not on the first, since the vowel of the second syllable is lower than that of the first. In (20e) the first two syllables are CVV. A stress falls on the second syllable in (20e) due to the fact that it has a lower vowel peak than that of the first syllable. The example in (20f) shows the same pattern except that it is the third

and fourth syllables that are of relevance. Crowhurst and Michael account for the effect of vowel quality on the determination of stress by proposing a vowel quality scale in addition to the coda scale. When all else is equal, the quality scale will determine stress going on to the syllable with a lower vowel peak.[17] To sum up the Nanti discussion from Crowhurst and Michael, Nanti only has two degrees of syllable weight: bimoraic syllables with long vowels and monomoraic syllables that can either be CV or CVC. However, other factors such as vowel quality and the presence of a coda consonant give the appearance of a system with more than two degrees of syllable weight. While a separate vowel quality scale can be motivated for other languages like Mordwin, discussed above, where quality determines stress placement, it remains to be seen whether coda consonants in multiple weight systems such as that in (17) for Hindi can be understood in terms of a separate coda scale and not as moraic weight. The division in Hindi in (17) where CVV and CVC pattern together seems more consistent with an analysis of three degrees of syllable weight as opposed to a coda scale as in Nanti.[18]

4 Syllable onsets and weight

One of the consistent findings in studies on syllable weight that is implicit in Trubetzkoy (1939) is that onset consonants do not play a role in the determination of syllable weight. (We delay the separate issue of initial geminate consonants until Section 5.) The onset is irrelevant to most processes that are sensitive to syllable weight. Consider the fairly common process of closed syllable shortening, where a long vowel shortens in a syllable closed by a coda (or moraic) consonant. In this process, a potential trimoraic CVVC syllable becomes bimoraic CVC. Now, if "heavy" or complex onsets contributed a mora to the weight of the syllable, we might expect to find vowel shortening processes where a potential trimoraic CCVV syllable would shorten to bimoraic CCV. However, to my knowledge, such cases of vowel shortening have not been reported in the literature. Further, minimal word effects do not ever seem to treat the onset consonant as moraic (again deferring the issue of initial geminate consonants until Section 5). For example, if a language requires prosodic words to be bimoraic, it does not distinguish between V, CV, and CCV potential words; all would be disallowed. Similarly, we do not find a minimal word pattern requiring that monosyllabic words begin with an onset whereas longer words can begin with a vowel.[19] The length or even the presence of the onset is irrelevant in meeting a minimal word requirement.

Similarly, in tone languages, tone is realized on moraic elements. This can include a coda consonant (in addition to vowels and syllabic consonants). However, tone in the phonemic or contrastive sense (e.g. lexical and morphological tonal melodies) never seems to have consistent realization on an onset, at least there are no clear instances of it (but see Topintzi 2006 for a possible example of a tonal onset in Kpelle). Furthermore, as noted in Section 1, processes of compensatory lengthening are typically not triggered by the loss of an onset consonant,

as in the Turkish example illustrated in (8a). This is consistent with Hayes's theory that when compensatory lengthening does occur, it involves the loss of a segment that was moraic. Potential counter-examples to this, such as the three cases noted by Kavitskaya (2002), all involve the deletion of a syllable initial sonorant consonant triggering vowel lengthening. As Hayes (1989a: 282) notes, such cases may involve insertion of a vowel before the deletion of the consonant, or perhaps one could suggest that these cases involve a sonorant being incorporated into the nucleus before it deletes.[20] In general, then, in languages that have processes sensitive to syllable weight, the onset never counts as moraic. For example, one can observe a wide range of mora sensitive processes described for Japanese including external phenomena such as speech errors and language games (Kubozono 1999; Tsujimura 2007) without encountering any phenomena where onset consonants contribute to syllable weight.

Nonetheless, there is one area in the literature regarding the possible participation of onsets in a weight-sensitive process: patterns of stress assignment. Do onset consonants ever play a role in stress processes, and, if so, does that provide evidence that they can be moraic or contribute weight to the syllable? We turn to this question, and suggest that onset consonants are never moraic.

Davis (1985) was probably the first to survey languages reported to have onset-sensitive stress and mentions about a dozen languages. Gordon's (2005) more recent survey includes thirteen languages reported to have onset-sensitive stress (though see Topintzi 2006 for critical comments on some of his examples as well as some additional cases). From these surveys as well as from other works such as Davis (1988) and Downing (1998), there seem to be only two types of onset-sensitive stress rules. The first type, which is more common, found amongst Australian languages (e.g. Arrernte) and Native American languages (e.g. Banawá), is where stress falls on the initial syllable if the word begins with a consonant and on the second syllable if the word begins with a vowel. The second type is where a phonological feature on an onset consonant is one of several factors in the determination of stress. (Topintzi (2006) considers a third type that combines the two types.) It is clear from these surveys that one never finds a language where stress falls on the syllables containing a complex (branching) onset, analogous to languages that place stress on heavy syllables. That is, there are no languages where CCV syllables are targeted for stress. One can maintain that the two types of onset-sensitive stress systems that do occur do not constitute cases of onset weight. First, consider the Arrernte type stress system where stress falls on the initial syllable if the word begins with a consonant and on the second syllable if the word begins with a vowel. While one may be tempted to suggest that in such languages the onset is moraic and stress falls on the leftmost bimoraic syllable, there are superior ways of analyzing such data without referencing moraicity. Downing (1998), for one, observes that onsetless syllables are exceptional to a range of different prosodic processes and can be considered as ill-formed syllables. She posits an analysis where there is a misalignment between morphological and prosodic constituents. Thus, a word-initial vowel would not be part of the prosodic word and outside the domain of stress. Other possible analyses

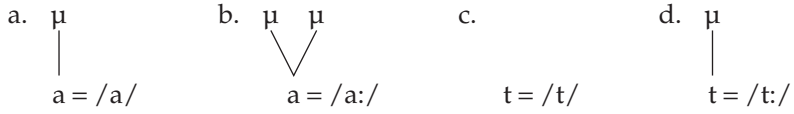
include having a constraint that aligns the left edge of a foot with a consonant (Goedemans 1996, 1998) or having a constraint that aligns the left edge of a stress syllable with a consonant (Topintzi 2006; Hyde 2007a). Any of these processes captures the stress pattern without having to posit a moraic onset and can be justified based on typological data. Thus, the most common type of onset-sensitive stress system that places stress on the initial syllable if the word begins with a consonant and on the second syllable if it begins with a vowel does not suggest that onset consonants are moraic.

With respect to languages in which some phonological feature of an onset consonant plays a role in the determination of stress, Davis (1988) posits that such cases involve rules of stress shift that occur after the stress has been assigned. That is, rules of stress shift (and stress deletion) may reference a feature on an onset. However, this would not necessarily entail that the presence of certain features on an onset consonant can make it moraic. On the other hand, Topintzi (2006) does make such a proposal. She observes that in three of the languages that have onset-sensitive stress, the South American languages Pirahã, Arabela, and Karo, stress can be attracted to a syllable with a voiceless consonant in its onset. In the well-known Pirahã case (Everett and Everett 1984; Everett 1988) primary stress falls on one of the last three syllables of the word, whichever one has a long vowel; if there is more than one with a long vowel or if there is none with a long vowel, stress falls on the rightmost one containing a voiceless consonant in its onset. In Arabela, primary stress on a final syllable will move to the penultimate syllable if the final syllable starts with a voiced consonant and the penultimate with a voiceless one. Finally, in Karo, a word-final stress will move to the penultimate syllable if the final syllable begins with a voiced obstruent and if the penultimate syllable begins with a voiceless obstruent or sonorant, as long as the final syllable does not have high tone, a nasalized vowel, or final sonorant consonant. Topintzi proposes that in all three of these languages voiceless consonants are moraic. Thus, they can attract stress. In support of this view, she notes a parallel with sonorant consonants being moraic in coda position. Sonorant consonants are the preferred coda type, so they can be moraic in coda position without obstruents being moraic. Similarly, voiceless obstruents are the preferred onset consonant type, so they can be moraic in the onset without treating other consonants as moraic in the onset. While this is an interesting proposal, all three of these languages can just as easily be analyzed as involving stress shift or with an onset prominence scale (or with a prominence projection referencing onsets as in Hayes's (1995a) analysis of Pirahã). Moreover, if voiceless consonants are moraic, we would expect to find languages where stress falls on the syllables in a word containing a voiceless onset, or we might predict a tendency for languages to avoid long vowels in syllables that begin with a voiceless obstruent. Given that such phenomena do not seem to occur, it may be preferable to view the apparent onset-sensitive nature of stress in Pirahã, Arabela, and Karo as involving stress shift or an onset prominence scale, and not reflecting moraic weight in the onset.[21] Consequently, it is still possible to maintain that (non-geminate) consonants in a syllable onset never contribute weight to the syllable.[22]

5 Geminate Weight

No issue in Hayes's (1989a) proposal regarding underlying moraic structure as in (5), repeated below as (21), has arguably generated as much controversy as his proposal in (21d) that geminate consonants differ from single consonants (21c) in that they are underlyingly moraic.[23]

(21) Underlying moraic representation (Hayes 1989a)



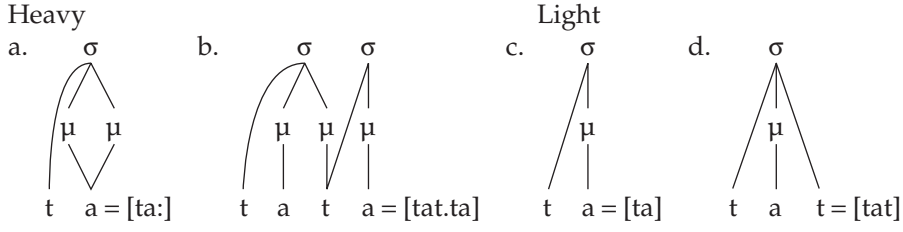
Hayes actually does not discuss the implications of the representation in (21d), but it is made clear in subsequent work by Selkirk (1990) and Tranel (1991), namely that there should be languages having the weight system shown in (22) in which a syllable closed by a geminate and a syllable with a long vowel act as heavy or bimoraic while a CV syllable and a syllable closed by a non-geminate consonant act as light or monomoraic (G = geminate consonant, C = non-geminate consonant).

(22) Syllable weight distinction based on geminates being underlyingly moraic:

Heavy	*Light*
CVV	CV
CVG	CVC

The system in (22) is predicted to occur under Hayes's theory in any language that allows long vowels and geminate consonants but in which Weight-by-Position does not apply. The moraic representation with syllable structure of (22) is given in (23).

(23) Surface syllabification of the division in (22)



As seen in (23b), a syllable closed by a geminate is bimoraic while one closed by a singleton consonant (23d) is monomoraic. Although there are many aspects of the geminate controversy that could be considered here, we will first focus our discussion on evidence concerning the division in (22). Specifically, are there

processes that treat CVG syllables as heavy while treating other CVC syllables as light? We will then discuss other evidence for the moraification of geminates in (21d), briefly reviewing evidence from morphology and the behavior of word-initial geminates.

As noted by Selkirk (1990) and Tranel (1991), if the syllable weight distinction in (22) exists as Hayes's theory predicts, we would expect to find languages where a syllable closed by a geminate (i.e. the first part of a geminate) acts as heavy while that closed by a non-geminate does not. Sherer (1994) and Davis (1994, 1999a, 2003) show from a variety of processes that there do seem to be languages that make such a weight division. We will take up evidence from closed syllable shortening and stress.[24] If we first consider closed syllable shortening, with the weight division in (22) one would expect to find a language where a long vowel shortens in a syllable closed by a geminate but not in one that is closed by a single consonant. Shortening would occur in the potential CVVG syllable in order to avoid a trimoraic syllable while shortening would not occur in CVVC since that would only be bimoraic. Kiparsky (2008c) mentions Swedish as a language where vowel shortening occurs before a geminate but not before a single coda consonant. Another language displaying this pattern of shortening is the Dravidian language Koya, discussed by Tyler (1969) and Sherer (1994) as well as by Davis (1999a), which the following discussion is based on. Koya has long vowels, coda consonants, and geminate consonants. There are words in Koya like those in (24a–c) where a long vowel can occur before a coda consonant. Crucially, as Tyler (1969: 6) observes, there are no words in which a long vowel occurs before a geminate. They are always short as in (24d). (All Koya data are cited from Tyler (1969) with the page numbers provided; vowel length is indicated by a colon; the transcription of the vowel quality is phonemicized and does not reflect the precise allophonic variant.)

(24) a. le:ŋga 'calf' (p. 11)
b. a:ɳɖa 'female' (p. 8)
c. ne:rs 'learn' (p. 76)
d. ett 'lift' (p. 76)

Moreover, cases are found where a stem-final long vowel shortens before a suffix beginning with a geminate, as in (25).

(25) a. ke: + tt + o:ɳɖu [kettoɳɖu] 'he told' (p. 39)
b. o: + tt + o:ɳɖu [otto:ɳɖu] 'he bought' (p. 38)

This shortening can be viewed as a way of avoiding trimoraic syllables. Shortening does not occur before a non-geminate consonant as the examples in (26) illustrate.

(26) a. na:l + ke [na:lke] 'tongue' (p. 47)
b. tuŋg + ana: + n + ki [tuŋgana:ŋki] 'for the doing' (p. 90)

In (26) a long vowel surfaces before a syllable-final singleton coda consonant. Since vowel shortening occurs before a geminate in (25), the Koya data in (24)–(26)

are consistent with the weight system in (22) in which CVV and CVG syllables are bimoraic but CVC syllables are light.[25]

While the above examples of Koya and Swedish are cases where vowel shortening occurs in syllables closed by a geminate, one can also find languages where vowel lengthening processes are prevented in CVG syllables but not in CVC syllables. This suggests that in such languages geminates are underlyingly moraic, though coda consonants in general are not; vowel lengthening then does not apply before a geminate since that would create a trimoraic syllable. A good example of this comes from Seto (Southeastern Estonian) discussed by Kiparsky (2008c). According to Kiparsky, this language has feet that are required to be trimoraic and this is normally implemented by foot-final vowel lengthening. Because of this process, a foot with the underlying sequence CV.CVC surfaces as CV.CVVC. However, given an input structure where the final consonant of the foot is part of a geminate, that is CV.CVG, no vowel lengthening occurs. This provides evidence that the geminate is underlyingly moraic; foot-final vowel lengthening need not occur in CV.CVG since the foot is already trimoraic.

A different case of a language that avoids the surfacing of CVVG syllables can be found in the West African language Fula as discussed by Paradis (1987) and Sherer (1994). Fula avoids CVVG syllables by degemination of the consonant. Importantly, as shown in (27), Fula allows CVVC syllables both morpheme-internally and over a morpheme boundary.

(27) CVVC syllables in Fula (Sherer 1994: 176)

a. kaakt-ɛ ‘spittle’ b. caak-ri ‘couscous’

Fula has a suffixation process that triggers the gemination of a root-final consonant. This is exhibited in the singular/plural alternations in (28). Because of a constraint in Fula requiring geminates to be [−continuant], a root-final continuant segment changes to a stop when it geminates. (I thank Abbie Hantgan for help on the Fula data.)

(28) Fula morphological gemination (Paradis 1987: 78)

	Stem (sg.)	*Suffixed form (pl.)*	*Gloss*
a.	lɛw	lebb-i	month
b.	lɛf	lepp-i	ribbon

Of relevance here is that when a long vowel precedes the stem-final consonant, gemination fails to occur, but the stem-final consonant nonetheless is realized as a stop. Consider the singular/plural alternations in (29).

(29) Lack of gemination after a long vowel (Paradis 1987: 80)

	Stem (sg.)	*Suffixed form (pl.)*	*Expected form*	*Gloss*
a.	laaw	laab-i	*laabb-i	road
b.	lɛɛs	lɛɛc-ɛ	*lɛɛcc-ɛ	bed

Given that gemination is part of this suffixing process, we note that the expected forms in (29), where the initial syllable would be CVVG, fail to surface as such. Rather, given the nature of the occurring suffixes forms in (29), it appears that degemination has occurred. This can be understood as the avoidance of a trimoraic CVVG syllable. Since CVVC syllables are allowed in Fula as seen in (27), Fula seems then to be a language that instantiates the weight system of (22) where CVG syllables are heavy but not other CVC syllables.

With respect to the stress evidence for the weight division in (22), probably the strongest case against the moraic analysis of geminate consonants is the observation from Tranel (1991) that there do not seem to be quantity-sensitive stress systems that support the weight division in (22) where stress would be attracted onto a syllable with a long vowel or closed by a geminate consonant but not on one closed by a non-geminate. Tranel (1991) points to the Uralic language Selkup as a language that has geminates but where CVG syllables are ignored by stress assignment even though syllables with long vowels attract stress. Consider the data below in (30). (The data in (30a–f) are given in Halle and Clements (1983: 189), while the data in (30g–h) are reported by Ringen and Vago (2011) for the Taz dialect of Selkup (citing the Selkup scholar, Eugene Helimski, p.c.), which has the same stress pattern as that shown by Halle and Clements.)

(30) Selkup (Halle and Clements 1983)

a.	qumó:qi	'two human beings'
b.	ú:cɨqo	'to work'
c.	u:cɔ́:mɨt	'we work'
d.	qúmɨnɨk	'human being (dat.)'
e.	ámɨrna	'eats'
f.	ú:cɨkkak	'I am working'
g.	ɛ́sükka	'(it) happens (occasionally)'
h.	ɛssí:qo	'to happen (already)'

In Selkup, primary stress falls on the rightmost syllable with a long vowel (30a–c) or on the initial syllable if there are no long vowels (30d). A CVC syllable does not count as heavy for stress (30e), even if the CVC syllable is closed by a geminate as seen in (30f–g). As noted by Tranel (1991), if stress is targeting bimoraic syllables and geminates are underlyingly moraic, then the second syllable in (30f–g) would be the rightmost bimoraic syllable. Both the vowel and the geminate would contribute a mora to the second syllable. The fact that (30f–g) do not receive stress on the second syllable seems to provide evidence against geminates being moraic.

The stress pattern of Selkup does not appear to be unique in ignoring geminate consonants. Davis (1999a: 41) points to the Altaic language Chuvash (Krueger 1961), which has an almost identical stress pattern to that of Selkup: stress is attracted to the rightmost syllable with a full vowel but CVG syllables are ignored.

Thus, in both Chuvash and Selkup, CVG syllables do not seem to function like bimoraic CVV syllables but instead act like monomoraic CV and CVC syllables.[26] Furthermore, languages where the stress pattern supports the syllable weight division in (22) seem rare. Davis (1994) discusses a Hindi dialect described by Gupta (1987) in which stress is attracted to the leftmost heaviest syllable in the word. The dialect treats both CVV and CVG syllables as bimoraic while CVC syllables behave as light. But as Curtis (2003) has pointed out, the pattern described by Gupta may be unusual among Hindi dialects in distinguishing CVG from CVC syllables.

One example of a language where stress assignment distinguishes CVG syllables from CVC syllables is the Uto-Aztecan language Cahuilla, and this is noted by Hayes (1995a). Consider the data in (31).

(31) Cahuilla stress (Hayes 1995a and references therein)

a.	tákalìčem	'one-eyed ones'
b.	čéxiwèn	'it is clear'
c.	táxmuʔàt	'song'
d.	qá:nkìčem	'palo verde (pl.)'
e.	héʔi kákawlà:qà	'his legs are bow-shaped'
f.	čéxxìwen	'it is very clear'

Following Hayes (1995a), the Cahuilla stress pattern in (31) reflects the assignment of moraic trochees starting from the left edge of the word. This is clearly shown in (31a–b). The comparison between (31c) and (31d) is interesting. In (31d), the first syllable is being treated as bimoraic given that there is a secondary stress on the second syllable: that is, in (31d) the initial syllable constitutes a moraic trochee on its own. There is a secondary stress on the second syllable since that would be the head of the second trochaic foot in the word. In (31c), even though the first syllable is closed, it is not treated as bimoraic. There is no secondary stress on the second syllable. Thus, the comparison between (31c) and (31d) shows that a syllable with a long vowel (i.e. the first syllable in (31d)) is regarded as bimoraic whereas a CVC syllable (i.e. the first syllable in (31c)) is regarded as monomoraic. This is also made clear by the form [kákawlà:qà] in (31e) where the CVC second syllable is skipped for stress, and the CVV third syllable forms a bimoraic foot on its own. The stress on the final syllable in (31e) indicates that Cahuilla allows a degenerate (i.e. monomoraic) foot in word-final position, as discussed by Hayes (1995a). Given this, it is noteworthy that the initial CVG syllable in (31f) counts as bimoraic forming a trochaic foot on its own. Just as in the case of (31d), there is a secondary stress on the second syllable in (31f). While Cahuilla provides evidence for the weight distinction in (22),[27] the general lack of stress evidence for (22) poses a challenge for the inherent underlying moraic analysis of geminates as in (21d). Perhaps, if more languages with the right set of properties are considered (i.e. stress languages with long vowels, coda consonants, and geminates), other cases supporting (22) will be found.[28]

We now briefly turn to evidence for the underlying moraification of geminates in (21d) that is independent of the weight system in (22). Based on Davis (1999a) there are cases where morphological allomorphy seems to be sensitive to the underlying moraification of the stem (as opposed to its surface moraic structure) and this can provide evidence for the underlying moraic nature of geminates. As one example, consider the Hausa plural pattern in (33)–(36) referred to as Class 3 plurals by Kraft and Kraft (1973) and discussed in such works as Newman (1972, 1992) and Leben (1980) (though the analysis to be presented here is somewhat different from these works). These plurals involve the suffixation of the two different allomorphs in (32) (where the C-slot in (32a) is realized as a consonant identical to the last root consonant).

(32) a. -aaCee b. -aayee

The data in (33) show nouns whose roots end in a single consonant while the data in (34) show nouns whose roots end in a consonant cluster. These roots select the plural allomorph in (32a). (As seen in the data, most singular nouns in Hausa end in a final long vowel extension that is not part of the root. Tones are not indicated in the data, but the affixation is accompanied by a HLH tone pattern over the whole plural form.)[29]

(33)		*Singular*	*Plural*	*Gloss*
	a.	dam-oo	dam-aamee	'monitor'
	b.	wur-ii	wur-aaree	'place'
	c.	kaf-aa	kaf-aafee	'small hole'

(34)				
	a.	gulb-ii	gul-aabee	'stream'
	b.	birn-ii	bir-aanee	'city'
	c.	kask-oo	kas-aakee	'bowl'

Now consider the data in (35) and (36) which select the plural allomorph in (32b). In this allomorph the suffix's second syllable has an onset realized by the default consonant [y], which Newman (1972) notes is found elsewhere in Hausa as an epenthetic consonant. (35) contains examples where the root vowel is long and (36) shows examples where the root vowel is a diphthong.

(35)		*Singular*	*Plural*	*Gloss*
	a.	zoom-oo	zoom-aayee	'hare'
	b.	kiif-ii	kiif-aayee	'fish'
	c.	suun-aa	suun-aayee	'name'

(36)		*Singular*	*Plural*	*Gloss*
	a.	ɓaun-aa	ɓaun-aayee	buffalo
	b.	mais-oo	mais-aayee	disused farm
	c.	gaul-aa	gaul-aayee	idiot

Based on the data in (33)–(36) one could generalize that if the root syllable contains a short vowel the allomorph -aaCee is selected, but if the root syllable contains a long vowel or diphthong then -aayee is selected. However, the data in (37) show that this generalization is not quite correct.

(37)		*Singular*	*Plural*	*Gloss*
	a.	tukk-uu	tukk-aayee	bird crest
	b.	tall-ee	tall-aayee	soup pot
	c.	gamm-oo	gamm-aayee	heat pad

The nouns in (37) have short root vowels, yet they nonetheless pattern like the nouns containing long vowels and diphthongs in (35) and (36) by taking the plural allomorph shown in (32b). The question that emerges then is what unifies the roots in (35)–(37) that distinguishes them from those in (33)–(34). The answer seems to reside in their underlying weight. The root forms in (35) and (36) have the shape CVV; those in (37a–c) have the shape CVG. These then are forms where the roots are at least bimoraic underlyingly, given Hayes's moraification algorithm as reflected in (21). On the other hand, the roots in (33) and (34) would be underlyingly monomoraic. We see then that the allomorph in (32a) only attaches to a noun root that is underlyingly monomoraic. The allomorph in (32b) attaches to noun roots that are at least bimoraic underlyingly. To be clear, the plural allomorphy, which is a lexical process in Hausa, is sensitive to the underlying mora structure of roots and not to the surface mora structure. This is because there is much evidence showing that surface CVC syllables are always bimoraic in Hausa. (Newman 1972 mentions vowel shortening in closed syllables, among other processes.) Though the initial CVC syllables in the singular forms in (34) would surface as bimoraic because of the weight of the coda consonant, the plural allomorphy, which is sensitive to the underlying moraic structure, treats (34) as monomoraic. This, then, can be seen as providing evidence for an underlying distinction between geminate consonants, which are moraic and singleton consonants, which are not.[30]

A final type of evidence for the moraic nature of geminates as in (21d), independent of the weight system in (22), comes from the behavior of word-initial geminates. Though such geminates are rare, they are attested in a number of languages. (In fact the dissertations of Muller (2001) and Topintzi (2006) are exclusively on initial geminates.) Muller (2001), whose study includes acoustic analyses of word-initial geminates, concludes that initial geminates are moraic in some languages but not in others. Topintzi (2006, 2008) focuses on languages where initial geminates pattern as moraic, and argues that such geminates constitute moraic onsets, thus providing a case where onsets carry weight.[31] An example of a language where a word-initial geminate patterns as moraic is Trukese (also called Chuukese). Consider the data in (38) and (39) (cited from Davis 1999b and Davis and Torretta 1998, and see references cited therein) that reflect a minimal word constraint on Trukese nouns. (Note that Trukese forms are given in transcription rather than in Trukese orthography.)

(38)

Underlying Representation	*Output Form*	*Gloss*	*Unattested output*
a. /maa/	[maa]	behavior	*ma
b. /təə/	[təə]	islet	*tə
c. /oo/	[oo]	omen	*o

(39)

Underlying Representation	*Output Form*	*Gloss*	*Suffixed form n = relational*
a. /etiruu/	[etiru]	coconut mat	[etiruu-n]
b. /ttoo/	[tto]	clam sp.	[ttoo-n]
c. /ččaa/	[čča]	blood	[ččaa-n]
d. /ssɔɔ/	[ssɔ]	thwart of a canoe	[ssɔɔ-n]

Trukese has a general process whereby a word-final long vowel shortens, as in (39). However, as (38) shows, shortening does not apply if the result would be monomoraic. This is because Trukese has a minimal word constraint that requires nouns to be bimoraic. The fact that the word-final vowel does shorten in (39b–d) strongly suggests that the initial geminate is moraic. That is, an output such as [tto] in (39b) is bimoraic with a mora being contributed by both the vowel and the geminate. As another example of a word-initial geminate acting as moraic, Topintzi (2006, 2008 and references cited therein) refers to stress evidence from Pattani Malay. In Pattani Malay there are no long vowels, and geminates only occur in word-initial position. Normally in Pattani Malay primary stress falls on the final syllable of a word, except when the word begins with a geminate consonant, in which case the initial syllable is stressed. This, too, can be taken as evidence for the moraification in (21d) where a geminate is underlyingly moraic.

Nonetheless, there are cases where word-initial geminates do not pattern as moraic. A good example is that of the Austronesian language Leti as discussed by Hume *et al.* (1997). Leti like Trukese has initial geminates and a bimoraic minimal word requirement, but unlike Trukese, Leti does not have words of the pattern shown in the output forms in (39b–d) consisting of an initial geminate followed by a short vowel. The lack of such words can be taken as strong evidence that the initial geminate does not count as moraic in Leti. Perhaps one can understand the difference between Trukese and Leti in terms of the language-specific phonotactics. In Trukese, word-initial geminates are permitted but there are no word-initial consonant clusters. In Leti on the other hand, not only are word-initial geminates permitted but almost any possible word-initial sequence of two consonants can occur with no apparent sonority restrictions between them. Given this patterning, one could analyze the first consonant of a word-initial cluster in Leti as being extraprosodic. The initial consonant of such a cluster is unrestricted and can be identical to the following consonant. This means that the word-initial geminate of Leti consists of a sequence of identical consonants; the first consonant of the sequence would be extraprosodic just like the first consonant of any other word-initial cluster. Still, the difference between Leti and Trukese suggests that there is no consistent behavior in the weight properties of word-initial geminates.[32]

Given this inconsistent behavior of geminate consonants, we suspect that the topic of the phonology of geminate consonants will continue to be a controversial one.

6 Weight Inconsistencies

So far in this chapter we have assumed that languages witness a weight consistency, such that, for example, if CVC syllables are heavy in a language (or if CVS syllables are heavy in a language), they consistently pattern as bimoraic throughout the language. Nonetheless, researchers have observed that weight inconsistencies not only occur but are quite common. These typically involve CVC syllables sometimes patterning as light and sometimes as heavy within the same language. In this section we will briefly examine two different situations where weight inconsistencies are found. In Section 6.2 we will consider what Hyman (1992) terms "moraic mismatches" or prosodic inconsistencies where within the same language certain syllable types act as heavy for one process but light for another. First, though, in Section 6.1, we will consider the somewhat different phenomenon of context dependent weight (Hayes 1994; Rosenthall and van der Hulst 1999) where the weight of a CVC syllable is based on its context within a word.

6.1 Context Dependent Weight

Context dependent weight is a phenomenon noted by such researchers as Kager (1989), Hayes (1994, 1995a), Alber (1997), Rosenthall and van der Hulst (1999), and Morén (1999, 2000) whereby CVC syllables surface as heavy (bimoraic) only in certain contexts. Specifically, one must consider the makeup of the other syllables within a word to know whether a particular CVC syllable will pattern as bimoraic.[33] Context dependent weight of CVC syllables occurs in certain stress systems whereby a CVV syllable in a word receives stress, but if a word has no CVV syllables, a CVC syllable receives stress. This exemplifies context dependent weight since a CVC syllable is heavy (bimoraic) only in words without long vowels. As an example of context dependent weight, both Rosenthall and van der Hulst (1999) and Morén (2000) discuss Kashmiri stress. Consider the Kashmiri data in (40).

(40) Kashmiri stress pattern (Rosenthall and van der Hulst 1999; Morén 2000)

a.	vah.ráa.vun	'to spread'
b.	báa.laa.dər	'balcony'
c.	yu.nu.vár.si.ti	'university'
d.	jəm.bɨr.zal	'narcissus'
e.	á.ni.ga.tɨ	'darkness'

The generalization illustrated by the stress patterns in (40) is that stress falls on the leftmost heavy syllable. While a CVV syllable is always bimoraic, a CVC

syllable can be heavy only in a word without long vowels (i.e. other bimoraic syllables). Concerning the data in (40), the items in (40a–b) show that in words containing long vowels, primary stress goes on the syllable containing the leftmost long vowel; (40c–d) indicate that if the word has no long vowels then primary stress goes on the leftmost closed syllable; otherwise, stress goes on the leftmost (initial) syllable as in (40e). (Note that the references cited for Kashmiri indicate that final syllables are never stressed and would be extrametrical.) The word-initial CVC syllable in (40a) does not pattern as bimoraic, while the word-initial CVC syllable in (40d) and the CVC syllable in (40c) do, since they receive stress. Thus, we see from the Kashmiri data that a CVC syllable can only surface as bimoraic in words lacking any CVV syllables.

In terms of a formal optimality-theoretic analysis (similar but not identical ones are offered by Rosenthall and van der Hulst 1999 and Morén (2000)), the Weight-to-Stress constraint forces a coda to surface as moraic in words lacking inherently bimoraic CVV syllables. In examples like (40c), then, the CVC syllable surfaces as bimoraic with stress since there are no other potentially bimoraic syllables to stress (given that the language does not lengthen underlying short vowels to make a syllable bimoraic). It is interesting that this analysis of context dependent weight can account for the occurrence of certain syllable weight hierarchies that were mentioned in Section 3. Given the Kashmiri stress data in (40), one could suggest that Kashmiri has a syllable weight hierarchy in which CVV syllables are heavier than CVC syllables which in turn are heavier than CV syllables. Stress would then fall on the leftmost heaviest (non-final) syllable in the word. However, given the contextual weight of CVC syllables, some of the cases of syllable weight hierarchies reported in the literature can be viewed as reflecting a system of context dependent weight. For example, the syllable strength hierarchy for Nanti in (19) is probably better analyzed as involving context dependent weight rather than as reflecting an independent coda strength hierarchy.[34]

It is important to mention another phenomenon that is considered by Rosenthall and van der Hulst (1999) to be a type of context dependent weight. Without doubt, the most common case of variable coda weight is final consonant extrametricality where in languages in which CVC syllables are normally bimoraic, a word-final CVC syllable functions as light. This is a pervasive phenomenon that occurs in many unrelated languages. (See Hayes (1995a: 58–60), in particular, for arguments supporting the notion of final extrametricality and Hyde (2007b) for an elaborated theory of non-finality.) In languages like Arabic, a CVC syllable normally attracts stress if it constitutes the penultimate syllable of the word but not if it constitutes the final syllable. In such languages, a penultimate CVC syllable would be bimoraic but a final CVC syllable would be monomoraic. In an optimality-theoretic analysis such as that of Rosenthall and van der Hulst (1999) the constraint ranking would normally select a coda that is moraic in the output, but a high-ranking non-finality constraint prevents a word-final coda from being moraic; thus a word-final CVC syllable patterns as light.[35]

A very intriguing observation regarding final extrametricality of CVC syllables put forward by Ham (1998) is that final CVC syllables are always extrametrical

in languages that have word-final geminates. This is because a word-final geminate is moraic and would need to be distinguished in final position from a potential moraic coda. Given the underlying representation of geminates as in (21d), final extrametricality of CVC syllables is able to preserve the contrast between an underlying final geminate and the corresponding final singleton consonant. The geminate of a final CVG syllable would surface as moraic while the singleton coda of the final CVC would be non-moraic. This difference is found in Arabic dialects where a final CVG syllable attracts stress, making it distinct (i.e. bimoraic) from a final CVC syllable which is light (monomoraic) and does not attract the stress. In a variety of other languages having word-final geminates examined by Ham (1998) the same distinction is made between final CVG and CVC syllables. If Ham's (1998) observation holds up to further scrutiny, it constitutes an interesting argument for the underlying moraification of geminate consonants. (See also Topintzi (2008: 175) for discussion on this point.)

6.2 *Moraic Mismatches*

Another type of weight inconsistency found in languages involves CVC syllables being treated differently depending on the specific process. One would expect that if CVC syllables act as bimoraic in a language they would act as bimoraic for all relevant weight-sensitive processes in that language. Nonetheless, researchers such as Crowhurst (1991), Steriade (1991), Hyman (1992a), Broselow (1995), Hayes (1995a) have noted what Hyman (1992a) refers to as moraic mismatches and others such as Fitzgerald (2012) call prosodic inconsistencies. This is the case where within a single language, CVC syllables sometimes act as heavy and sometimes as light, depending on the process at issue. Hyman (1992a) gives a variety of examples from Bantu languages. Consider the data Hyman (1992a: 258–259) provides from the Runyambo-Haya dialect cluster of Tanzania in (41). (High tone is represented by an acute accent.)

(41) Runyambo-Haya dialect cluster (Tanzania)

a. Assign a high tone to the second mora of a verb stem
   1. ni-tu-rim-á 'we are cultivating'
   2. ni-tu-siíg-a 'we are smearing'
   3. ni-tu-jend-á 'we are going'

b. Reduplication (the final a vowel reduplicates as long if the stem syllable is monomoraic, but it reduplicates as short if the stem syllable is bimoraic)
   1. (ku-) lim-a → (ku-) limaa-lima 'to cultivate'
   2. (ku-) siig-a → (ku-) siiga-siiga 'to smear'
   3. (ku-) genda → (ku-) genda-genda 'to go' *(ku-) gendaa-genda

In (41a) we see a process whereby a high tone is assigned to the second mora of a verb stem (underlined in 41a). As shown in the third example in (41a), a preconsonantal nasal consonant does not add a mora to the verb stem. In (41b) we

see a verbal reduplication process whereby the final vowel /a/ reduplicates as long if the verb stem is monomoraic but as short if the stem syllable is bimoraic. Here, the third example in (41b) shows that the preconsonantal nasal of the verb stem does add a mora. Thus, according to Hyman (1992a), Runyambo-Haya exemplifies a moraic mismatch. Mora count is different for tone (41a) versus reduplication (41b). Tone assignment treats the preconsonantal nasal as non-moraic, while the same nasal adds a mora to the verb stem with respect to reduplication.[36]

Along similar lines, Steriade (1991) notes that in Khalkha Mongolian the stress rule treats CVC syllables as light or monomoraic since stress is attracted to a CVV (bimoraic) syllable skipping over CVC syllables, but a bimoraic constraint on minimal verb stems regards both CVV and CVC as heavy (bimoraic), given that no CV stems are found but CVV and CVC stems occur. Steriade (1991) suggests that it is only processes like stress and tone that are involved in moraic mismatches treating otherwise bimoraic CVC syllables as light.[37] Specifically, according to Steriade (1991), languages may restrict tone and stress bearing elements to those above a certain sonority threshold where pitch realization would be clearest, and this does not necessarily mean that CVC syllables are monomoraic.[38]

The notion of moraic mismatches, that certain weight-sensitive processes in a language may consider CVC syllables as light while others treat them as heavy, is most fully developed in a series of important works by Gordon (1999, 2001, 2002, 2004b). Gordon argues for a radical departure viewing weight as process-driven rather than language-driven. In Gordon's view, moraic mismatches should be the expected case. It just depends on the process whether certain CVC syllables will be treated as heavy, not on the language. In order to show this, Gordon (2004b) reports on a survey of six weight-sensitive phenomena in approximately 400 languages. Phenomena that he surveyed included stress, tone, poetic metrics, compensatory lengthening, minimal word requirements, and templatic restrictions. As an example of his findings supporting his view, he found that most languages treated syllables closed by an obstruent (CVO) as heavy with respect to minimal word constraints but as light with respect to tonal phenomena. Another finding is that syllables closed by a sonorant, (CVS), frequently pattern as heavy with CVV syllables for tonal phenomena, but this patterning is rare for stress. An example of this, also discussed by Blevins (2004), is Lhasa Tibetan, which has both tone and stress. Lhasa contour tones, which are only realized on heavy syllables, treat CVV and CVS as heavy, but CVO as light. On the other hand, Lhasa stress, which treats CVV syllables as heavy, considers all CVC syllables as light, including CVS syllables. Lhasa Tibetan represents a typical finding in Gordon's survey. Generally, with respect to stress patterns, Gordon found that languages either treated just CVV syllables as heavy (as in Lhasa Tibetan) or both CVV and CVC syllables as heavy; it was rare for a language to treat CVS as heavy with respect to stress without also treating CVO as heavy. But, as mentioned, CVS syllables often patterned with CVV syllables as heavy with respect to tonal phenomena.

Gordon maintains that the reason why tone and stress pick out different syllable types as heavy is because they have a different phonetic basis. Gordon (2004:

285–286) explains the patterning of CVV and CVS as heavy for tone along the following lines:

> The physical correlate of tone is fundamental frequency, which is only present in voiced segments . . . Crucially, the fundamental frequency profile of a segment or syllable (and hence its tonal profile) is cued not only by the fundamental itself but also by the higher harmonics . . . The presence of harmonics greatly enhances the presence of fundamental frequency . . . [T]he more crucial harmonics for the perception of the fundamental, the low frequency harmonics (House 1990), are typically present in sonorants . . . In contrast to sonorants, obstruents provide either minimal or no cues to fundamental frequency.

Thus, tone is sensitive to the presence of certain harmonics found in sonorants but not in obstruents. Consequently tonal phenomena can treat both CVV and CVS syllables as heavy. On the other hand, according to Gordon, stress is sensitive to the overall auditory energy in the syllable rhyme. Sonorancy and voicing are two of the best features for predicting higher energy values in the rhyme. Gordon (2004b) makes the interesting claim that a major factor in determining whether or not CVC syllables act as heavy for stress is the nature of the language's coda inventory. For example, if we compare Khalkha Mongolian, which treats CVC syllables as light for stress, with Finnish, which treats them as heavy, Gordon observes that in Khalkha Mongolian there are more voiceless consonants that can occur in coda position than voiced ones (sonorants included) while in Finnish there are more voiced consonants in coda position than voiceless ones. Consequently, in general, Finnish codas have more auditory energy than Khalkha codas and so Finnish CVC syllables behave as heavy with respect to stress while Khalkha Mongolian CVC syllables behave as light. Thus, not only does Gordon argue for the process-specific nature of syllable weight, he argues further that the processes do not share a single phonetic basis for what makes a syllable heavy.

While Gordon's research offers a new and different perspective on syllable weight, a perspective that focuses on process types rather than on issues that emerge from a moraic theory of quantity as has been the focus of this chapter, critical assessment of Gordon's work can be found in Curtis (2003), Topintzi (2006), and de Jong (2000). Curtis (2003: 290) notes that all of Gordon's mismatches involve CVC syllables with varying codas and do not involve CVV syllables. She concludes that his findings support phenomenon-specific weight-by-position rules that can reference features such as sonorant, but they do not challenge the inherent weight of vowels or even of geminates, nor the structural representation of syllable weight. Topintzi (2006) focuses her criticism on stress being related to overall auditory energy noting, among a variety of potential problems, the lack of sonorant onsets affecting stress in languages where onsets seem to matter for stress. As was mentioned in Section 4, Topintzi observes that when a feature of an onset consonant influences stress it is usually a voiceless consonant that attracts the stress. This is the opposite of what Gordon's theory of auditory energy would predict. An important criticism of Gordon's theory connecting the patterning of

CVC syllables with respect to stress with overall auditory energy in the syllable rhyme comes from Ahn's (2000) typological survey of stress systems, which is discussed in detail by de Jong (2000). Ahn makes a strong claim, based on stress descriptions of 136 languages, that quantity-sensitive unbounded stress systems (such as that of Khalkha Mongolian) always treat CVC syllables as light while bounded systems such as Finnish can treat them as heavy. This difference between how CVC syllables are treated in bounded vs. unbounded systems does not fall out from Gordon's theory. De Jong (2000) suggests that the underpinnings of stress are different in the two systems: bounded systems can reflect delimitative intonation marking over syllables while unbounded systems reflect loudness or prominence. (See de Jong (2000) for details as to why this can result into a different treatment of CVC syllables.)

7 Conclusion

In this chapter we have surveyed a variety of issues that emerge from the moraic representation of quantity. Our focus has been almost exclusively on phonological issues concerning the patterning of possible moraic elements and their role in determining syllable weight. We have spent less time examining alternative representational proposals regarding syllable weight (though some of these are mentioned in Section 1 and in the notes), not only because of space limitations but also because recent critical works (e.g. Curtis 2003; Kraehenmann 2001; Ringen and Vago 2011) assume familiarity with moraic theory.

We have dealt with the phonetics of syllable weight only briefly here, in the discussion regarding Gordon's work in Section 6.2. One might expect that if CVV and CVC syllables stand out for a variety of processes it is because they are longer than CV syllables. But a direct correlation with duration was already observed to be problematic by Trubetzkoy (1939), who notes phonetic contextual effects on duration that are not moraic, such as contextual effects on vowel duration depending on the nature of surrounding elements. A good example showing that phonetic duration does not necessarily translate to syllable weight is the observation by Chen (1970) that vowels are longer before voiced consonants. This seems to have no effect on stress placement. We do not find stress systems where primary stress falls on a syllable containing a vowel before a voiced consonant.[39] Perhaps for such reasons, Hayes (1995a: 271) emphasizes phonological duration rather than phonetic duration in his discussion of weight: "weight can be thought of as a property of the time dimension: a syllable is heavy because it is long. This is the viewpoint of moraic theory: the moras form an abstract characterization of a syllable's phonological duration." But even a view of phonological duration was called into question by Newman (1972: 320) who maintained that, ". . . the distinction between heavy and light syllables cannot be assumed on a priori grounds to be phonologically analyzable in terms of units of duration nor to be phonetically correlated with actual-time differences." Gordon (2004b) also takes a position against a strictly durational characterization of syllable weight, noting that it is

problematic in languages that have both CVV and CVC syllables but where only CVV syllables pattern as heavy, since he finds in his durational study of Khalkha Mongolian (which treats only CVV as heavy with respect to stress) that the duration of rhymes in both CVV and CVC syllables are distinct from rhyme duration of CV syllables. Thus, he concludes that duration alone is not a good fit for determining the basis of syllable weight in such languages. What emerges from phonetic explorations into syllable weight like that of Gordon as well as de Jong (2000) is that there is no one phonetic correlate of syllable weight. Thus, one could maintain that syllable weight is an abstraction and that moras form its abstract characterization.

ACKNOWLEDGMENTS

This chapter could not have been written without the work done by the participants in the syllable weight seminar at Indiana University in Spring 2007. In particular, I thank Majdi Bamakhramah, Nick Henriksen, Chisato Kojima, Dongmyung Lee, Paco Montaño, Jung-yueh Tu, Junghyoe Yoon, and Kenji Yoshida. I also thank Steve Chin, Chris Green, Iskra Iskrova, Brian José, Jungsun Kim, Traci Nagle, Nathan Sanders, and Natsuko Tsujimura for giving me detailed comments on a preliminary version of this chapter. Further, I thank Brett Baker, Ken de Jong, Daniel Dinnsen, Tracy Alan Hall, Abbie Hantgan, Michael Marlo, Paul Newman, Anne Pycha, Robert Vago, and Islam Youssef for discussion on specific aspects of this paper. Finally, I thank the editors, especially John Goldsmith, for their valuable comments in improving the quality, if not the quantity, of this chapter.

NOTES

1 In this chapter, length is transcribed either by a colon or by a sequence of identical letters. For example, geminate-p will be transcribed either as [p:] or [pp].

2 Note, though, that such contextual phonetic differences may eventually become relevant phonologically as in the development of compensatory lengthening in Friulan, as discussed by Kavistskaya (2002).

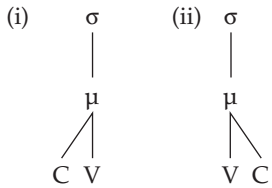
3 This is not to say that issues of representation were unimportant in the discussion of quantity in the pre-generative period. Relevant discussion can be found in Trubetzkoy (1939: 173–174) of the English translation, and in Hockett (1955: 76–77). Somewhat relatedly, with respect to English there is a longstanding debate on whether the tense-lax vowel distinction is really one of quantity or quality and how the difference should be represented. Duponceau (1818) is critical of those who do not recognize a quantity distinction in the description of English vowel sounds (pp. 239–240) and further suggests that the English length distinction should be represented by a diacritic on a short vowel. Trager and Bloch (1941) took vowel quantity in English as *h* since length and *h* can be interpreted as being in complementary distribution. On the other hand, Chomsky and Halle (1968) viewed the English vowel contrast as one of quality (tense vs. lax) rather than quantity.

4 All references to Trubetzkoy 1939 in this chapter are to the 1969 English translation of it. See Anderson (1985: 100–106) for discussion regarding Trubetzkoy's view of quantity.

5 According to Hock (1986b: 90) a notion like the mora can be traced back to the Sanskrit Grammarians of the fifth century B.C.E. The earliest linguistic reference to mora mentioned by the *OED* is from volume XI (p. 591) of the first edition of the *Encyclopedia Americana* (1832), where the term is equated with a short syllable. Bloomfield (1933: 110) refers, to mora as "an arbitrary unit of relative duration." Important works within the framework of generative phonology that incorporate or develop a theory of mora include Newman (1972), Hyman (1985), Hock (1986a), and McCarthy and Prince (1986). These all slightly differ from one another and from Hayes (1989a). It is Hayes's theory that is most influential in current work within moraic phonology.

6 We use the terms "onset," "nucleus," and "coda" to make reference to the different positions within the syllable, with onset being the syllable initial consonant or consonants, the nucleus being the vowel or peak of the syllable, and coda being the syllable-final consonant or consonants. Here, we make no assumption regarding the formal status of these as constituents of the syllable, but see Davis (2006) for an overview.

7 In this chapter it is assumed that a non-moraic consonant, whether an onset or a coda, attaches directly to the syllable node. While this is a fairly standard assumption for onset consonants and was assumed by Hayes (1989a), Katada (1990) argues that an onset shares a mora with the following vowel as in (i). With respect to the coda consonant, Broselow *et al.* (1997) give phonetic arguments for a non-moraic coda sharing a mora with a preceding vowel as in (ii):



See the works cited for discussion of these technical issues.

8 In this chapter we make the common assumption that an intervocalic geminate is heterosyllabic as shown in (7e) and that an intervocalic consonant cluster syllabifies as heterosyllabic as well (ignoring cases where such clusters are also possible complex onsets). Any intervocalic consonant is normally assumed to syllabify as an onset, so that /VCV/ would syllabify as [V.CV], but see Blevins (2004) for further discussion.

9 Kavitskaya (2002) notes three cases where the deletion of an onset consonant leads to compensatory lengthening. This is unexpected in Hayes's theory since onsets are not typically moraic. It may be of relevance that each of these cases involves the deletion of a sonorant in the onset. See the discussion in Note 20 regarding compensatory lengthening in Samothraki Greek, one of the cases discussed by Kavitskaya.

10 Throughout this paper, unless otherwise noted, when syllable types such as CV, CVV, and CVC are listed or discussed, it should be understood that the initial C is to be interpreted as standing for C_0 (meaning zero or more initial consonants). This reflects the general weightlessness of syllable onsets, an issue that will be discussed in Section 4.

11 The moraic theory elaborated on here predicts that there are no languages in which CVC patterns as heavy while CVV patterns as light. This is because long vowels are

underlyingly bimoraic whereas a coda consonant is not underlyingly moraic; so a syllable with a long vowel would always be treated as bimoraic. As noted by Blevins (1995: 237), this prediction seems correct. A separate issue is whether there are languages that lack CVV syllables but where CVC syllables are heavy. Hayes (1989a: 290) mentions Ilokano and Spanish as possible examples and Hayes (1995a: 205) cites the Cariban language Hixkaryana. Seneca (Prince 1983) may be another example, given that its accent system is sensitive to closed syllables but lacks long vowels phonemically. However, Trubetzkoy's (1939) discussion of syllable weight and Zec's (1988, 1995b) proposal on sonority thresholds for moraic segments (to be discussed in Section 2) predicts that the presence of bimoraic CVC syllables in a language implies the presence of CVV syllables in that language. We will not be considering this issue further here.

12 While the system in (13) has [+sonorant] as a minimal sonority threshold on what can be moraic, the system in (10a) would have no minimal sonority threshold on what can be moraic (so even obstruents can be moraic); the system in (10b) would have a strict sonority threshold in which only [+sonorant], [–consonantal] elements (i.e. vocalic phonemes) can be moraic.

13 There are a variety of syllable structure issues concerning diphthongs that go beyond the scope of this chapter. These include whether the glide element of a diphthong is part of a syllable nucleus or syllable margin and whether rising diphthongs (i.e. a sequence of an on-glide and a vocalic peak) pattern as monomoraic or bimoraic. See Davis and Hammond (1995), Baertsch (2002), Smith (2003), and Levi (2004) for relevant discussion.

14 While both Tiv and Lithuanian allow only higher sonority consonants to be moraic, they differ with respect to the surfacing of lower sonority consonants such as voiceless stops in coda position. In Lithuanian they surface as non-moraic in coda position, but in Tiv they never surface in codas. This suggests that Weight-by-Position (i.e. the constraint or rule that requires codas to be moraic) still plays a role in languages where there is a minimal sonority threshold on what can be moraic. Weight-by-Position can be violated in Lithuanian but not in Tiv.

15 There are languages that witness an interesting interaction of sonority constraints on moraicity and the Coda Condition. In Ponapean (Goodman 1997) only sonorant codas are permitted but they must obey the Coda Condition (ignoring some loanwords). Thus, in Ponapean, the only codas permitted are sonorant consonants that share place features with the following onset and these behave as moraic. We can analyze Ponapean as having, minimal sonority threshold on moraic elements requiring them to be sonorant, along with the Coda Condition as an edge constraint. In Campidanian Sardinian (Davis and Baertsch 2005, and references cited therein), at least in initial syllables which have less restricted phonotactics than other syllables, a coda consonant must be either a high sonority rhotic (e.g. [ar.ba] 'white') or a consonant that obeys the Coda Condition, such as a nasal homorganic to a following onset or the first part of a geminate. The effect of this is that a consonant of lower sonority in the coda must obey the Coda Condition whereas a high sonority coda (i.e. rhotic) need not obey it. See Davis and Baertsch (2005) for a partial optimality-theoretic analysis and discussion. Further, Basbøll (2005) offers an analysis of Danish in which both obstruents and sonorants appear in the coda position, but only sonorants are moraic; this frequently coocurs with the stød.

16 An alternative way of understanding the hierarchy in (19) is to consider it a case of context dependent weight along the lines of Rosenthall and van der Hulst (1999). On such a view, CVC syllables are normally monomoraic but can be bimoraic in a

context where there would otherwise be only monomoraic syllables in the word. This is discussed in Section 6.1.

17 The various scales that Crowhurst and Michael (2005) employ (such as the vowel quality scale) to help determine stress location can be seen as a type of prominence projection along the lines of Hayes (1995a). Hayes employs a prominence projection for atypical situations where factors other than syllable weight, such as tonal quality or certain features on phonemes, play a role in stress.

18 One issue not discussed here is whether a segment can ever display a three-way contrast in quantity. Trubetzkoy (1939: 180–181) doubted such cases existed phonologically even in languages like Estonian and Lapp (Saami), suggesting that the appearance of multiple degrees of quantity were a phonetic effect resulting from other factors in these two languages; however, see Bye, Toivonen, and Sagulin (2008) for evidence of a three-way phonemic consonantal length contrast in Inari Saami.

19 Odden (2006) discusses a minimality pattern in the Bantu language Zinza where bisyllabic words must either begin with a consonant or a long vowel, but never a short vowel. Superficially, this may look like a case where a CV (initial) syllable patterns together with a syllable beginning with a long vowel, perhaps suggesting the bimoraicity of CV syllables. Odden, though, accounts for the Zinza pattern by referencing a high ranked constraint militating against prosodic words beginning with short vowels. The potential moraicity of onsets is not at issue.

20 The clearest case in the literature where the deletion of an onset consonant triggers compensatory lengthening involves /r/ deletion in Samothraki Greek where a word-initial /r/ and /r/ as a second member of an onset cluster deletes triggering compensatory lengthening of a following vowel; /r/ does not delete in coda position, but does delete intervocalically without triggering compensatory lengthening. Hayes (1989a: 283) suggested an analysis involving epenthesis, but as Topintzi (2006) shows, such an analysis is not motivated synchronically. Topintzi presents a thorough discussion of the Samothraki data and analyzes it as an instance where the deletion of a non-moraic segment results in the addition of a mora by compensatory lengthening, thus constituting a clear counter-example to Hayes's (1989a) theory. A different way of looking at the Samothraki Greek data is to posit, abstractly, that /r/ deletes when it is forced into the nucleus of the syllable; thus it is moraic when it deletes. This is suggested by Kiparsky (2011). Key to such an analysis is that /r/ can never surface as an onset. Kavitskaya (2002) provides a similar view suggesting that the /r/ is vocalic enough to be interpreted as additional vowel length.

21 One can speculate that it is the higher pitch that occurs on vowels after voiceless consonants that affects stress placement in Arabela and Karo. This could be the case in Karo where stress is retained on a final syllable that has a high tone even if it begins with a voiced consonant and the penultimate begins with a voiceless one. Hyman (2006b), though, notes that pitch does not seem to be a factor in the determination of stress in Pirahã and that the phonetic length of a voiceless consonant may indeed be making a contribution in the determination of stress.

22 Perhaps a perceptual basis for the weightlessness of syllable onsets can be found in Goedeman's (1998) experimental study where, using CVC synthetic stimuli, he observed that listeners were less sensitive to duration fluctuations in the onset consonant than in either the vowel or coda consonant. However, as Goedemans notes, this may be an effect of onset weightlessness; that is, the weightlessness of onsets influences listeners' ability to perceive durational contrast in the onset. See Gordon (2005) for discussion on Goedemans' work.

23 This controversy is reflected in certain important response articles that argued against this underlying moraic representation of geminates, such as Selkirk (1990), which argued for a two root node theory of geminates and Tranel's (1991) article, which posited a principle of equal weight for codas whereby geminates were moraic or non-moraic depending on the patterning of other codas in the language. More recently, Ringen and Vago (2011) have argued for a universal segmental length representation of geminates as in (i) with no inherent weight properties:

(i)

Ringen and Vago maintain that all evidence for a single root analysis of geminates as illustrated in (21d) is reanalyzeable with structure like that in (i) and that there are phenomena that are only compatible with that in (i). A recent defense of the moraic theory of geminates can be found in Topintzi (2008). Other researchers have proposed composite representations for geminates such as Schmidt (1992) and Hume *et al.* (1997) who incorporate both an X-tier and a moraic tier in their analysis of geminates. Specifically, geminates are represented as a single phoneme linked to two X-slots and could be moraic in coda position if other codas in the language are moraic. Moreover, the controversy over geminates has fostered a number of dissertations which have a focus on the phonology of geminates. Some of the more important ones include Curtis (2003) who proposes a model of moraic representation that combines the two root node theory of geminates with moraic theory, Ham (1998), Keer (1999), Kraehenmann (2001), Morén (1999), Muller (2001), Sherer (1994), and Topintzi (2006). It is impossible in this chapter to discuss the wide variety of interesting issues and proposals that are raised in these dissertations.

24 See Davis (1994, 1999a) for the range of processes considered to bear on the weight division in (22) and see Curtis (2003) and Ringen and Vago (2011) for critiques.

25 Curtis (2003: 169–170) suggests that the lack of word-internal CVVG syllables in Koya may be due to a shortening effect that geminate consonants have on preceding vowels since the perceptual cues for vowel length can be blurred in CVVG syllables; thus, Curtis maintains that vowel shortening before geminates is independent of the issue of the moraic status of geminates.

26 Proponents of the moraic theory of geminates have suggested various strategies for such languages whereby, geminate consonants do not seem to pattern as moraic. Topintzi (2008), who for the most part maintains the underlying moraic view of geminates, suggests that weightless geminates in a language like Selkup are represented by double consonants with two root nodes rather than as a single root node linked to a mora like (21d). Davis (2003) suggests that the stress pattern of languages like Selkup do not necessarily argue against the moraic theory of geminates, rather, from an optimality-theoretic perspective, the pattern can be a consequence of certain high-ranking stress constraints that have the effect of ignoring the bimoraicity of any CVC syllable. See Davis (2003: 95) for elaboration.

27 Hayes (1995a) notes that Cahuilla CVC syllables closed by glottal stops are also treated as bimoraic, though not other CVC syllables. In this light, it is worth recalling the Ngalakgan data in (15) which constitutes an almost opposite pattern to Cahuilla in closed syllables whereby syllables closed by geminates and those closed by a glottal stop are skipped over for stress and thus act as monomoraic while other CVC syllables

(at least those where the coda has its own place features) are treated as bimoraic. In order to maintain the moraic theory of geminates, given the Ngalakgan stress data, one would need to maintain that, underlyingly, Ngalakgan geminates are indeed moraic; that is they have the underlying moraic structure in (21d); they just do not surface as moraic because of a constraint that requires moraic elements to have their own place features. In this way, Cahuilla geminates would reflect their underlying moraic structure while Ngalakgan geminates do not. With respect to the issue in Cahuilla of a syllable closed by a glottal stop being treated as heavy, it may be possible to view this as reflecting a sonority threshold condition on what can be moraic. It is well-known that glottal type consonants can either pattern as highly sonorous elements or as obstruents though this has not been much discussed in the literature, but see Churma and Shi (1995) and Parker (2002). This variable behavior can be understood as to whether a language treats the laryngeal articulator as a place of articulation in the vocal tract or as just reflecting a particular state of the glottis. In the case of the latter, the laryngeal consonant would be a sonorant since there would be a free flow of air in the vocal tract. If this is the case for Cahuilla, then it would provide an interesting case where moraic codas are either highly sonorous consonants (glottal stops) or a geminate. Such a combination of moraic codas can be understood through Morén's (1999, 2000) distinction between coerced weight and distinctive weight. See Note 33 for discussion on this distinction.

28 A very interesting case of a language in which CVV and CVG syllables pattern together with respect to stress is San'ani (Yemen) Arabic as described by Watson (2002: 81–82) who specifically notes their patterning together as opposed to CVC syllables. In this language, stress falls on the rightmost non-final CVV or CVG syllable in the word. If there are no such syllables (and ignoring a possible final superheavy syllable), then stress falls on the rightmost non-final CVC syllable up to the antepenultimate syllable; otherwise, it falls on the leftmost CV syllable. Thus, in words where there is a (non-final) CVC syllable and a (non-final) CVV or CVG syllable, it is either the CVV or CVG syllable that receives stress. The priority of CVV and CVG syllables in this language implies that CVC only acts as heavy in words in which there are no underlyingly bimoraic syllables (CVV or CVG). That is, Weight-by-Position could only apply in a word that would otherwise have no bimoraic syllables. The patterning of CVG with CVV is best understood here if the geminate is underlyingly moraic. The underlying moraicity of geminates in San'ani Arabic is further supported by Watson's (2002: 82) observation that CVC syllables are never stressed in pre-antepenultimate position whereas both CVV and CVG syllables can be. See Watson (2002) for an analysis that incorporates the view that geminates are underlyingly moraic. We suspect that syllables closed by geminates are special (and distinct from CVC syllables) with respect to weight properties in other dialects of Arabic as well, but this topic has yet to be fully explored.

29 The Hausa data in this section have been discussed with Paul Newman, the leading authority on the Hausa language (e.g. Newman (2000), though he disagrees with the analysis suggested here that the allomorph selection between (32a) and (32b) reflects an underlying weight distinction.

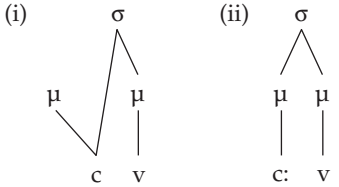
30 The Hausa plural pattern is made somewhat more complicated by the data in (i) that involve roots ending in a nasal followed by a homorganic stop.

(i)

	Singular	*Plural*	*Gloss*
a.	kund-ii	kund-aayee	'notebook'
b.	gunt-uu	gunt-aayee	'stub'

Davis (1999a) argues that these "partial geminates" can be considered underlyingly moraic and so can pattern with the roots in (37) with a geminate. (Davis (1995b) surveyed the weight behavior of partial geminates, showing that depending on the language, they could pattern like geminates.) Newman (1992) argues for an analysis of (i) where the nasal is incorporated into the nucleus and so they would pattern with other roots having complex nuclei like in (35) and (36). Newman (electronic communication, September 3, 2007) further suggests that the different behavior of CVCCVV singulars (e.g. 34 and 37) with respect to the plural allomorphy is due to the segmental characteristics of the abutting consonants rather than to underlying moraicitiy or other metrical factors.

31 Following a suggestion in Hayes (1995a), Davis (1999b) proposes that word-initial geminates are moraic but that the mora is not part of the syllable onset. His representation is in (i) while Topintzi's moraic onset representation is given in (ii) (where the vowel of the syllable is also shown).



One difference between (i) and (ii) is that (ii) predicts that onset geminates could occur word-internally, not just at the beginning of the word. In support of (ii) Topintzi (2008) provides interesting evidence from Marshallese that word-internal geminates are syllabified as onsets and are not heterosyllabic (7e) as commonly assumed in moraic theory.

32 José and Auger (2005) show that within a single language not all initial geminates pattern the same. According to them, in Vimeu Picard (phrase-)initial geminates differ in their representation as to whether they have a single set of features or two sets of identical features linked to two root nodes. Based on phonological patterning they argue that initial [ll] has the former representation while initial [nn] has the latter.

33 Of relevance to this topic is an important distinction that Morén (1999, 2000) makes between distinctive weight and coerced weight. Distinctive weight refers to underlying moraic structure that is reflected on the surface; geminate consonants, for example, would have distinctive weight. Coerced weight occurs when a non-moraic input segment surfaces as moraic. Context dependent weight can be seen as a type of coerced weight. However, coerced weight is more extensive than context dependent weight. For example, consider a language where Weight-by-Position always applies so that codas consistently surface as moraic. Such a language would have coerced weight since a non-moraic input consonant surfaces as moraic (in coda position), but it would not be an instance of context dependent weight since coda weight is not dependent on the makeup of the other syllables in the word. Morén makes an important observation that distinctive weight is not subject to sonority restrictions while coerced weight is. Thus, while Weight-by-Position may be restricted to sonorant consonants as discussed in Section 2, there are no universal sonority restrictions on distinctive weight; one does not find implications such that the presence of an obstruent geminate implies the presence of a sonorant geminate. This seems to be a correct observation about geminates as noted by Morén (1999), Blevins (2004), and implicit in Trubetzkoy's

(1939) discussion of geminates, though see Kawahara (2007) concerning certain tendencies regarding which consonant types are more likely to be geminate.

34 Klamath is another language cited in the literature (e.g. Hayes 1995a; Blevins 2006b) that has a weight hierarchy in which CVV syllables are heavier than CVC syllables which in turn are heavier than CV syllables. Stress falls on the rightmost heavy syllable in Klamath (with final CV and CVC syllables being extrametrical), but CVC syllables only count as heavy in words without long vowels. This can be analyzed in terms of contextual weight, as in Rosenthall and van der Hulst (1999). See also Yu's (2005) analysis of Washo stress and reduplication that makes crucial use of contextual weight of CVC syllables.

35 The specific nature of final extrametricality varies among languages. For example, in Cairene Arabic it is just the word-final consonant that is extrametrical while in Latin it is the entire final syllable. Also languages vary as to whether and how they incorporate final extrametrical elements into higher prosodic structure. For pertinent discussion regarding the analysis of English and German word-final syllables see Hall (2002).

36 More detailed discussion and analysis of the behavior and representation of preconsonantal nasals in Bantu languages can be found in Hyman and Ngunga (1997) and Downing (2005).

37 A somewhat similar mismatch to that in Khalkha Mongolian can be found in the Uto-Aztecan language Tohono O'odham. Fitzgerald (2012) documents that in Tohono O'odham CVC syllables behave as consistently bimoraic for the prosodic morphology of the language that includes processes of reduplication and gemination, but such syllables do not inherently attract stress. The language has a quantity-insensitive trochaic stress system that assigns alternating stress from the left edge of the word ignoring the apparent bimoraic nature of CVC syllables. CVV syllables in Tohono O'odham are restricted to word-initial position (at least in the native vocabulary), which is the location of primary stress.

38 The occurrence of moraic mismatches brings up certain representational issues regarding moraic structure that are discussed by Broselow (1995), which I will not discuss here, other than to mention Hayes's (1995a: 300) proposal of a moraic grid whereby sonorous (i.e. vocalic) moras would have two levels of grid marks and less sonorous moras would have one level. Certain processes like Haya tone or Khalkha Mongolian stress would make reference to the second level of grid marks, while other processes such as minimal stem or word constraints would make reference to the lower level of grid marks. There are other ways of analyzing these mismatches in a constraint-based approach to phonology, which we do not discuss here.

39 This suggests that lengthening due to phonetic factors is different from length that is phonological or distinctive. An interesting study that shows such a difference is Pycha (2007, 2009). She observes that Hungarian has two lengthening processes: a phrase-final process that lengthens a segment immediately adjacent to the phrase boundary and a morpho-phonological lengthening process whereby certain suffixes trigger gemination of a stem-final consonant. Pycha shows that when an affricate is targeted to be lengthened, the two processes implement the lengthening in different ways even though the overall duration is essentially the same. Specifically, with phrase-final lengthening, which Pycha considers to be phonetic, the lengthening of the targeted affricate occurs mainly in the fricative part of the affricate, but in the morpho-phonological lengthening triggered by the suffix the lengthening mainly occurs in the closure part of the affricate.

5 Stress Systems

MATTHEW GORDON

1 Introduction

Stress refers to increased prominence associated with a certain syllable or syllables in a prosodic domain. The study of stress is complicated by the existence of considerable cross-linguistic variation in the acoustic correlates of stress, the domain over which stress is assigned, the presence of secondary stress, and the relationship between stress and other types of prominence, such as phrasal pitch accents (see Beckman 1986; Hayes 1995; Ladd 1996; and Gussenhoven 2004, 2007 for overviews of these issues). Nevertheless, the formal investigation of stress has been a fruitful area of research in the phonology literature since the seminal work on generative metrical stress theory in the 1970s and early 1980s (e.g. Howard 1972; Liberman and Prince 1977; Hayes 1980; 1995; Prince 1983, 1990; Halle and Vergnaud 1987; Halle and Idsardi 1995). The last 30 years have witnessed important advances in both the typological knowledge of stress and its formal analysis, though many of the basic observations about the data serving as the basis for early generative work still underlie current research in metrical stress theory.

The advent of the constraint-based framework of Optimality Theory (OT) has provided a new framework for analyzing stress systems (e.g. Crowhurst and Hewitt 1994; Baković 1996, 1998; Walker 1996; Alber 1997, 2005; Eisner 1997; Kenstowicz 1997; Elenbaas 1999; Elenbaas and Kager 1999; Kager 1999, 2001, 2007; Gordon 2002a; Hyde 2002; McCarthy 2003b). One feature of OT that has been a boon to metrical stress theory is the relative ease with which analyses may be computationally implemented. This has facilitated evaluation of the predictive power of metrical theories and has also enabled testing of learning algorithms that model the acquisition of a stress system by language learners (e.g. Tesar and Smolensky 2000; Tesar 2004, 2007; Hayes and Wilson 2008).

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

This chapter examines the current state of metrical stress theory both from a typological and theoretical perspective. The structure of the chapter is as follows. Section 2 provides a typological overview of the various types of quantity-insensitive stress systems found cross-linguistically, focusing on word-level stress. Section 3 sketches a representative foot-based approach to these data couched within Optimality Theory. Section 4 introduces syllable weight effects observed in stress systems, while Section 5 discusses the relationship between word-level stress and phrase-level prominence. Section 6 examines the relative merits of foot-based and grid-based theories of stress in terms of their typological coverage. Finally, Section 7 summarizes the chapter.

2 Typology of quantity-insensitive stress systems

Quantity-insensitive stress systems, that is, those in which syllable weight is not relevant in conditioning stress placement, can be broadly divided into two groups based on whether stress rhythmically falls on syllables at regularly spaced intervals within a word or whether it is fixed on a syllable at or near one or both edges of a word. The Australian language Maranungku (Tryon 1970) provides a representative example of a rhythmic stress pattern. Primary stress in Maranungku falls on the first syllable of a word and secondary stress docks on odd-numbered syllables after the first one. Note that I use the IPA symbols for primary ˈ and secondary ˌ stress in the examples throughout the chapter.

(1) Stress in Maranugku (Tryon 1970)

ˈtiralk	‘saliva’
ˈmæræˌpæt	‘beard’
ˈjaŋarˌmata	‘the Pleiades’
ˈŋaltiˌritiˌri	‘tongue’

A variant of this pattern is found in another Australian language, Pintupi (Hansen and Hansen 1969, 1978), in which stress falls on odd-numbered syllables but does not fall on final syllables. The result is a sequence of unstressed syllables, a stress “lapse,” at the end of odd parity words.

(2) Pintupi stress (Hansen and Hansen 1969)

ˈtʲutaja	‘many’
ˈpuɭiŋˌkalatʲu	‘we (sat) on the hill’
ˈtʲamuˌlimpaˌtʲuŋku	‘our relation’
ˈʈiɭiˌriŋuˌlampatʲu	‘the fire for our benefit flared up’
ˈkuraˌnʲuluˌlimpaˌtʲuɻa	‘the first one (who is) our relation’

In some languages, stress falls on even-numbered syllables rather than odd-numbered ones. For example, primary stress falls on the second syllable in Osage (Altschuler 2009) and secondary stress docks on even-numbered syllables after the second one.

(3) Osage stress (Altschuler 2009)

ɑːˈleː	‘I left’
nãːˈxo	‘break by foot’
^{h}paːʃˈt͡seka	‘strawberry’
ðyːˈ^{h}kaːmã	‘to ring the bell’
xoːˈt͡soðiːbˌrã	‘smoke cedar’
ãːˈwãlaːˌxyɣe	‘I crunch up my own (e.g. prey) with teeth’

Macedonian (Lunt 1952; Franks 1987) is a language with a fixed stress pattern. A single stress in Macedonian falls on the antepenultimate syllable of a word and there are no reported secondary stresses.

(4) Macedonian stress (Franks 1987)

ˈzborot	‘word (def. sg.)’
ˈdonesi	‘bring (2nd sg. imper.)’
voˈdenitʃar	‘miller’
vodeˈnitʃari	‘miller (pl.)’
vodeniˈtʃarite	‘miller (def. pl.)’

The typology of rhythmic and fixed stress systems can be classified into three sub-types as follows. The first of these is the strict binary pattern involving stress on every other syllable, as in Maranunku. There are four logically possible types of strict binary systems if one varies the edge of the word at which the alternating pattern originates and whether the alternating stress pattern begins with a stressed syllable, that is, “peak-first,” or unstressed syllable, that is, “trough-first”: stress on odd-numbered syllables counting from the left, stress on even-numbered syllables counting from the right, stress on odd-numbered syllables starting at the right edge, and stress on even-numbered syllables commencing at the left edge. These four possibilities are shown schematically in (5).

(5) Typology of pure binary stress systems

Pattern	*Schematic forms*	*Example Lgs.*
1. Odd-numbered from L to R	ˈσσˌσσˌσ, ˈσσˌσσˌσσ	Czech (Kučera 1961), Maranungku (Tryon 1970)
2. Even-numbered from R to L	σˌσσˈσσ, ˌσσˌσσˈσσ	Cavineña (Key 1968), Warao (Osborn 1966)
3. Even-numbered from L to R	σˈσσˌσσ, σˈσσˌσσˌσ	Araucanian (Echeverría and Contreras 1965), Sirenikski (Menovshchikov 1975)
4. Odd-numbered from R to L	ˌσσˌσσˈσ, σˌσσˌσσˈσ	Chulupí (Stell 1972), Urubú Kaapor (Kakumasu 1986)

There are also a few languages that employ a ternary stress system, stressing every third rather than every other syllable. For example, Cayuvava (Key 1961, 1967) stresses every third syllable counting from the right edge of a word.

(6) Stress in Cayuvava (Key 1961, 1967)

'eɲe	'tail'
'ʃakahe	'stomach'
ki'hibere	'I ran'
ari.'u.utʃa	'he came already'
ˌdʒihira'ri.ama	'I must do'
maˌrahaha.'e.iki	'their blankets'
ikiˌtapare'repeha	'the water is clean'
ˌtʃa.adiˌroboβu'rurutʃe	'ninety-nine (first digit)'
meˌdarutʃeˌtʃe.iro'hi.iɲe	'fifteen each (second digit)'

Turning to the fixed stress systems, there are five docking sites for stress in these languages: the first syllable (e.g. Chitimacha: Swadesh 1946), the last syllable (e.g. Atayal: Egerod 1966), the penultimate (second-to-last) syllable (e.g. Albanian: Hetzer 1978), the antepenultimate (third-to-last) syllable (e.g. Macedonian: Lunt 1952; Franks 1987), and the peninitial (second) syllable (e.g. Koryak: Zhukova 1972). In most languages, only one of these syllables receives stress. However, there are also languages in which two of these syllables are stressed in a single word. In these systems, termed "hammock" (Elenbaas and Kager 1999) or "dual" (Gordon 2002a) stress patterns, there is one stress at or near the right edge and the other at or near the left edge. For example, Lower Sorbian (Janas 1984) places primary stress on the first syllable and secondary stress on the penultimate syllable. The secondary stress is suspended in trisyllabic words in order to avoid a sequence of adjacent stresses, a stress "clash."

(7) Lower Sorbian stress (Janas 1984)

'pisasʲ	'write'
'dɔbrɨ	'good'
'wɔsʲtsɔjska	'fatherland'
'psʲijasʲɛl	'friend'
'spewaˌjutsɨ	'singing'
'dɔpredˌkarski	'progressive'

There are also hybrid systems in which stress falls rhythmically on alternating syllables but also occurs on a fixed syllable at the opposite edge of the word from which the rhythmic pattern originates. For example, stress in the South Conchucos variety of Quechua spoken in Peru (Hintz 2006) falls on the penultimate syllable and on even-numbered syllables counting backwards from the penultimate. In addition, the initial syllable is stressed. There is some variation between

discourse data and elicited data as to which of the stresses is the strongest. In elicited forms, the stress on the penultimate is judged by speakers to be the primary stress, whereas in discourse data the stress on the initial syllable is primary. Forms illustrating stress in South Conchucos Quechua appear in (8), with the location of the primary stress reflecting discourse pronunciations.

(8) South Conchucos stress (Hintz 2006)

ˈʃumaq	‘pretty’
ˈimaˌkuna	‘things’
ˈʧupanˌkimanˌɬachi̥	‘you would likely have just gotten drunk’
ˈʧakranˌtsikːuˌnataˌraːʧir	‘yet our gardens supposedly’
ˈpiˌtapis	‘anybody’
ˈtuˌʃukuˌnaqḁ	‘dancers’
ˈwaˌraːkaˌmunqaˌnatʃi̥	‘hopefully it will appear at dawn’

Indonesian (Cohn 1989) displays a minor deviation from this pattern in words with an odd number of syllables, where the alternating pattern is suspended where it would result in a stress clash, that is, a sequence of adjacent stressed syllables. Thus, unlike South Conchucos Quechua which has the following stress pattern in a word of seven syllables, ˌσˌσσˌσσˈσσ, an Indonesian word with the same number of syllables would lack the stress on the second syllable, that is, ˌσσσˌσσˈσσ. Systems like the one found in South Conchucos Quechua in which rhythmic stress is observed even in clash contexts may be termed “binary plus clash” patterns, in contrast to “binary plus lapse” systems like that of Indonesian, in which a rhythmically placed stress fails to appear where it would clash with an adjacent fixed stress (Gordon 2002).

3 The Formal Analysis of Stress

The fundamental contribution of generative metrical stress theory is its formal treatment of stress as a prominence relation holding between syllables. Certain syllables are metrically strong, characteristically reflected in their attraction of stress, while others are metrically weak and thus reject stress. Theories differ, however, in how they represent this relative prominence. The most common approach is to assume that words can be broken down into smaller constituents called “feet” (Hayes 1980; Selkirk 1980b), which consist of a single stressed syllable and typically one unstressed syllable. A representative foot-based theory of stress is presented in Hayes (1995a). In Hayes’ theory, the Maranungku word ˈjaŋarˌmata would be represented as in (9).

(9) Word level (x . . .)
Foot level (x .)(x .)
ˈja ŋar ˌma ta

The first two syllables are grouped into a foot, while the last two syllables form a separate foot. Feet of the Maranungku type are "trochaic," that is, the stressed syllable precedes the unstressed syllable, in contrast to "iambic" feet, in which the stressed syllable follows the unstressed syllable, as illustrated by the Osage word xoːˈt͡soðiːbˌrɑ̃ 'smoke cedar' (10).

(10) Word level (. x . .)
Foot level (. x)(. x)
xoː ˈt͡so ðiː bˌrɑ̃ 'smoke cedar'

The higher tier of constituency is the word level, where the first stressed syllable is the metrically strongest one in the word in both Maranungku and Osage, as reflected in the word-level grid mark that it receives.

3.1 A Foot-based Metrical Stress Theory

In this section, we examine the typological coverage provided by a metrical stress theory employing feet. The analysis considered here is Kager's (2007) Optimality-theoretic approach, although the metrical representations assumed in his account can be translated easily into a rule-based framework like that of Hayes (1995). We now briefly introduce the constraints assumed in Kager's analysis.

First, two constraints are relevant in determining whether feet are trochaic or iambic in a language: FtType=Trochee and FtType=Iamb. Parse-Syl requires that syllables be parsed into feet. Another constraint, Ft-Bin, requires that feet be binary either at the syllabic or moraic level. The determination of the relevant prosodic level of analysis at which Ft-Bin applies depends on the role of quantity (syllable weight) in a language's stress system. In languages like those considered thus far with quantity-insensitive stress, Ft-Bin applies at the level of the syllable, whereas in languages with quantity-sensitive stress, Ft-Bin is relevant at the moraic level (for more on quantity-sensitive stress, see Section 4). If Ft-Bin is ranked above Parse-Syl, the result is a binary stress system with a prohibition against monosyllabic, degenerate, feet. The Pintupi stress system in which stress falls on odd-numbered non-final syllables instantiates a binary system banning monosyllabic feet, for example, (ˈtʲamu)(ˌlimpa)(ˌtʲuŋku) 'our relation', (ˈtiˌli)(ˌriŋu)(ˌlampa)tʲu 'the fire for our benefit flared up'. If the opposite ranking obtains, strict binarity is produced even in cases where a monosyllabic foot results, as in Maranungku, for example, (ˈjaŋar)(ˌmata) 'the Pleiades', (ˈŋalti)(ˌriti)(ˌri) 'tongue'.

The directionality of footing is determined by two alignment constraints: All-Ft-R and All-Ft-L, which require that all feet fall at the right and left edge, respectively, of a prosodic word. The simplest case is a stress system with a single stress per word, a pattern that results from the ranking of one of the All-Ft-X constraints above Parse-Syl. Depending on the type of foot and which foot-alignment constraint is highly ranked, different single stress systems are produced. Initial stress and penultimate stress reflect trochaic feet, where initial stress, that is, (ˈσσ)σσσ, results from the ranking All-Ft-L >> Parse-Syl while

penultimate stress, that is, σσσ('σσ), is attributed to the ranking All-Ft-R >> Parse-Syl. Peninitial stress and final stress are both the result of iambic feet with peninitial stress resulting from a left-aligned foot, that is, (σ'σ)σσσ and final stress reflecting a right-aligned foot, that is, σσσ(σ'σ). The final type of single stress system to account for, antepenultimate stress (e.g. Macedonian), results from nearly the same rankings as those accounting for penultimate stress, except that a ban on parsing the final syllable into a foot, Nonfinality, outranks All-Ft-R; the result is a trochee spanning the antepenultimate and penultimate syllables, that is, σσ('σσ)σ.

The promotion of one of the stressed syllables to primary stress is a function of two alignment constraints requiring, in the case of Align-Head-L, that the prosodic word begins with the primary stress foot or, in the case of Align-Head-R, that the prosodic word end in the primary stress foot.

If the relative ranking of Ft-Bin and Parse-Syl is varied along with the parameters of directionality and foot type, eight possible patterns are generated. These eight patterns, seven of which are attested, are depicted schematically in Table 5.1.

In addition to the four pure binary patterns (a–d in the table), there are four systems in which deviations from binarity are observed in words with an odd number of syllables. Two of these systems (e and f) have stress lapses at the word periphery in words with an odd number of syllables, whereas two have stress clashes. If Ft-Bin outranks Parse-Syl, the Pintupi-type pattern with stress on odd-numbered syllables from the left minus the final syllable and its mirror-image pattern of stress on odd-numbered syllables from the right minus the

Table 5.1 Eight stress patterns generated by permuting the ranking of Ft-Bin, Parse-Syl, All-Ft-X, FtType=Trochee and FtType=Iamb.

	Stress pattern	*Attested?*	*Example language*
a.	('σσ)('σσ)('σ),('σσ)('σσ)('σσ)	Yes	Maranungku (Tryon 1970)
b.	(σ'σ)(σ'σ)σ,(σ'σ)(σ'σ)(σ'σ)	Yes	Osage (Altshuler 2009)
c.	σ('σσ)('σσ),('σσ)('σσ)('σσ)	Yes	Cavineña (Key 1968)
d.	('σ)(σ'σ)(σ'σ),(σ'σ)(σ'σ)(σ'σ)	Yes	Urubú Kaapor (Kakumasu 1986)
e.	('σσ)('σσ)σ,('σσ)('σσ)('σσ)	Yes	Pintupi (Hansen and Hansen 1969; 1978)
f.	σ(σ'σ)(σ'σ),(σ'σ)(σ'σ)(σ'σ)	No	
g.	(σ'σ)(σ'σ)('σ),(σ'σ)(σ'σ)(σ'σ)	Yes	Ojibwa (Kaye 1973; Piggott 1980)
h.	('σ)('σσ)('σσ),('σσ)('σσ)('σσ)	Yes	South Conchucos Quechua (Hintz 2006)

initial syllable are produced. The former is the result of leftward alignment of feet, while the latter reflects rightward alignment of feet. The latter pattern appears to be unattested, in keeping with the general absence of evidence for initial stress avoidance in languages. The two binary systems with stress clashes in odd parity words (g and h) that are generated turn out to be attested. If Ft-Bin is ranked below Parse-Syl and feet are trochaic, stress falls on even-numbered syllables from the right plus the initial syllable, the pattern found in South Conchucos Quechua. If Ft-Bin is ranked below Parse-Syl and feet are iambic, stress will fall on even-numbered syllables plus the final syllable, which corresponds to the stress system of Ojibwa (Kaye 1973; Piggott 1980).

The generation of "hammock" stress systems with two stresses per word requires a pair of alignment constraints that ensure that word edges align with foot edges. These alignment constraints thus differ from All-Ft-L and All-Ft-R, which require that foot edges align with word edges. Align-PrWd-Left mandates that a word begins with a foot and Align-PrWd-Right requires a word to end with a foot. Align-PrWd-Left and Align-PrWd-Right perform an important role in generating hammock stress systems. The initial plus penultimate stress pattern found in languages like Lower Sorbian reflects the existence of a trochaic foot at each edge of the word, that is, ('σσ)σσσ('σσ). Both Align-PrWd-Left and Align-PrWd-Right are satisfied at the expense of All-Ft-L and All-Ft-R. In trisyllabic words, the ranking of FtBin over Align-PrWd-Right ensures that only a single foot is formed at the beginning of the word: ('σσ)σ not *('σσ)('σ) or *('σ)('σσ).

The Align-PrWd-X constraints also play a crucial role in generating binary systems with a stress lapse word-internally in words of a certain shape. To illustrate this, consider the analysis of the stress system of Garawa (Furby 1974), in which stress falls on even-numbered syllables counting from right to left but skips over an even-numbered peninitial syllable in favor of the initial syllable. The result is a stress lapse following the initial syllable (which carries main stress) in words with an odd number of syllables. Examples of Garawa stress appear in (11).

(11) Garawa stress (Furby 1974)

ˈjami	'eye'
ˈpunʲala	'white'
ˈwatʲimˌpaŋu	'armpit'
ˈnariŋinˌmukuˌnʲinaˌmira	'at your own many'

Let us consider the analysis of this pattern is some detail since it requires a rather complex set of rankings to produce: Align-PrWd-Left >> Parse-Syl >> All-Ft-R >> All-Ft-L. In addition to not violating Align-PrWd-Left, the winning candidates also satisfy Ft-Bin and FtForm=Trochee. A sample tableau illustrating the analysis of stress in Garawa appears in (12). Violations of the foot-alignment constraints committed by each foot are given as numerals separated by a comma, where the total number of violations of a constraint is the sum of the violations committed by each foot.

(12)

punʲala	Ft-Bin	FtFrm=Trochee	Align-PrWd-Left	Parse-Syl	All-Ft-R	All-Ft-L
☞ ('punʲa)la				*	1	
pu('nʲala)			*!	*		1
(pu'nʲa)la		*!		*	1	
('punʲa)(ˌla)	*!				0,1	0,2
('pu)(ˌnʲala)	*!				0,2	0,1
nariŋinmukunʲinamira	Ft-Bin	FtFrm=Trochee	Align-PrWd-Left	Parse-Syl	All-Ft-R	All-Ft-L
☞ ('nari)ŋin(ˌmuku)(ˌnʲina)(ˌmira)				*	0,2,4,7	0,3,5,7
('nari)ŋin(ˌmuku)nʲi(ˌnami)ra				**!*	1,4,7	0,3,6
na('riŋin)(ˌmuku)(ˌnʲina)(ˌmira)			*!	*	0,2,4,6	1,3,5,7
('nari)(ˌŋinmu)(ˌkunʲi)na(ˌmira)				*	0,3,5,7!	2,4,7

To account for ternary stress, Elenbaas and Kager (1999) employ the constraint *Long-Lapse, which requires an unstressed syllable to be adjacent to either a stressed syllable or the word edge. This constraint has the effect of banning a sequence of three consecutive unstressed syllables word-internally or two at a word edge. Ternary stress systems result from the ranking *Long-Lapse >> All-Ft-X >> Parse-Syl. As we have seen in the analysis of single stress systems, All-Ft-X has the effect of minimizing the number of feet in a word. However, *Long-Lapse ensures that there is no more than one unparsed syllable intervening between feet. For example, feet in Cayuvava (stress on every third syllable from the right) are trochaic and separated by a single syllable: (ˌtʃa.a).di.(ˌro.bo).βu.('ru.ru).tʃe 'ninety-nine (first digit)'.

3.2 Constraining the Foot-based Theory: Position Specific Rhythmic Constraints

Kager (2001) is an attempt to constrain the overgeneration of unattested patterns committed by the traditional foot-based metrical stress theory. In order to preclude

the unattested mirror image of Pintupi, stress on odd-numbered syllables counting from the right minus the initial syllable, Kager posits that local rhythmic constraints working in conjunction with the constraints requiring alignment of word edges with feet (Align-PrWd-Left and Align-PrWd-Right) govern the directionality of footing rather than the All-Ft-X alignment constraints. For example, a constraint Lapse-At-End requires that stress lapses be confined to the right edge of a word, penalizing forms with stress lapses in non-final position. This position specific anti-lapse constraint ensures that lapses are preferred word-finally to other positions, including word-initially. Thus, the candidate with a right-to-left iambic parse and an initial stress lapse cannot under any ranking emerge victorious over a competing candidate with a foot at the left edge (13) (from Kager 2007: 219).

(13)

σσσσσσσ	*Lapse	Lapse-at-End	Align-PrWd-L	Align-PrWd-R
☞ (σ'σ)σ(σ'σ)(σ'σ)	*	*		
σ(σ'σ)(σ'σ)(σ'σ)	*	*	*!	

Kager (2001) also notices another skewing in the typology. In most binary plus lapse systems involving an internal lapse, the lapse occurs immediately adjacent to the primary stressed foot. Thus, in Garawa, which stresses even-numbered syllables from the right and places primary stress on the initial syllable, the stress lapse follows the initial foot in odd parity words: ('σσ)σ(ˌσσ)(ˌσσ). In Piro (Matteson 1965), which stresses odd-numbered syllables counted from the left and places primary stress on the penultimate syllable, the lapse precedes the final foot: (ˌσσ)(ˌσσ)σ('σσ). Kager proposes that these systems result form a constraint requiring that lapses occur adjacent to a primary stressed foot: Lapse-at-Peak. This constraint ensures that lapses adjacent to the main stressed foot will be preferred over those in other positions.

Kager also finds that stress clashes in binary plus clash stress systems involve two secondary stresses rather than a primary plus a secondary stress. For example, in South Conchucos Quechua, the stress clash in odd parity words involves the initial and peninitial syllables, both of which have secondary stress, at least in words produced in isolation. Furthermore, stress clashes in binary plus clash systems invariably involve a foot at the word periphery, as in South Conchucos Quechua. To account for these two observations, Kager posits two constraints: *Clash-at-Peak bans stress clashes involving a primary stress and Clash-at-Edge requires a clash be restricted to a word edge.

A prediction made by this approach is that there should not be any binary plus internal lapse systems where the lapse is not adjacent to the primary stressed foot. Apparent cases where the lapse does not occur adjacent to the primary stressed foot must thus be reanalyzed in different terms. Such cases involve initial

dactyls where primary stress falls on the penultimate syllable and secondary stress is on even-numbered syllables counting from the right plus the initial syllable, for example, (ˌσσ)σ(ˌσσ)(ˈσσ). Alber (2005) argues that most of these cases (Polish, German, Indonesian) of initial dactyls involve loan words where the stress patterns of the donor language have been preserved.

4 Quantity-sensitive Stress

In many languages, an alternating binary stress pattern is interrupted by certain "heavy" syllables, which attract stress themselves even if they are immediately adjacent to a stressed syllable (see Davis this volume for more on syllable weight). After the heavy syllable, the normal binary stress count resumes. For example, stress in Chickasaw (Munro and Ulrich 1984; Munro and Willmond 1994; Gordon 2004a) falls on even-numbered syllables counting from the left and on heavy syllables, which are CVC and CVV in Chickasaw. In addition, the final syllable is stressed as well. (Primary stress falls on the rightmost long vowel, otherwise on the final syllable in words without long vowels.)

(14) Chickasaw stress

ˌisso'ba	'horse'
ˌtʃon'kaʃ	'heart'
ˌbaʃ'po	'knife'
a'boːkoˌʃiʔ	'river'
'baːˌtamˌbiʔ	'name'
'ʃiːˌki	'buzzard'
tʃaˌlak'kiʔ	'Cherokee'
ˌokˌfok'kol	'type of snail'
'naːɬtoˌkaʔ	'policeman'
'aːˌtʃomˌpaʔ	'store'

The attraction of stress by heavy syllables in quantity-sensitive stress systems means that several stressed syllables may occur in succession, thereby creating multiple violations of *Clash, as in words like ˌokˌfok'kol and 'baːˌtamˌbiʔ. An additional constraint is thus necessary to account for the stress-attracting property of heavy syllables.

4.1 Quantity-sensitive Stress in a Foot-based Framework

In a foot-based theory, quantity-sensitivity is typically incorporated directly into the foot type constraint. Hayes (1995) suggests that all iambic stress systems are quantity-sensitive and proposes three basic foot types. Some languages parse syllables into quantity-insensitive trochees, while quantity-sensitive stress systems may be either trochaic or iambic. In quantity-sensitive systems, the binary stress

count is calculated at the level of the mora rather than the syllable. Heavy syllables consist of two moras, while light syllables have one mora. Feet are minimally bimoraic, consisting of one heavy syllable or two light syllables. Thus, the Chickasaw word ˌisso'ba 'horse' would have the moraic and foot structure in (15), where moras are indicated by the Greek letter μ.

(15) (x)(. x)
ˌi$_{\mu}$ s$_{\mu}$ so$_{\mu}$'ba$_{\mu}$

4.2 Quantity-sensitivity in Non-binary Stress Systems

Quantity-sensitivity also plays a role in a number of languages with non-binary stress systems. For example, many single stress languages display quantity-sensitivity. Thus, in Yana (Sapir and Swadesh 1960), stress falls on the initial syllable in words with only light (CV) syllables (16a), but on the leftmost heavy syllable (CVV or CVC) if any are present (16b).

(16) Yana stress (Sapir and Swadesh 1960)

a.	'p'udiwi	'women'
b.	si'bumk'ai	'sandstone'
	su'k'oːniyaː	'name of Indian tribe'
	tsini'jaː	'no'

Single stress systems displaying quantity-sensitivity, sometimes called "unbounded" stress systems (Hayes 1995), like that of Yana have been treated in different ways in the theoretical literature. One approach assumes that all heavy syllables in unbounded stress systems carry at least secondary stress, which can be difficult to hear, and that an edgemost stress is promoted to primary stress by ER-L or ER-R (McCarthy 2003b). For example, the Yana word si'bumˌk'ai 'sandstone' would have two stresses, of which the leftmost one is primary. In words consisting of only light syllables, Align-PrWd-Left would position stress on the initial syllable. Under this approach, heavy syllables are assumed to constitute monosyllabic feet and words with only light syllables are treated as having a disyllabic foot at the word edge feet (Baković 1998; McCarthy 2003b). One potential drawback to this analysis, however, is its assumptions about secondary stress, which would ideally find support from instrumental studies and/or other phonological diagnostics. An alternative that still assumes a single stress per word is to posit a constraint against heavy syllables in pretonic syllables, in the case of leftward orientation, or posttonic syllables, in the case of rightward orientation.

4.3 Scalar Quantity-sensitivity

In some languages, quantity-sensitivity is scalar such that more than two degrees of weight are distinguished, with stress seeking out the heaviest syllable in the

hierarchy (see Davis this volume for more on scalar weight distinctions). In the simpler type of these scalar weight systems, all degrees of weight are operative within the same window. For example, in the variety of Hindi described by Kelkar (1968) there are three degrees of weight: superheavy (CVVC and CVCC), heavy (CVV and CVC), and light (CV). Stress falls on the heaviest syllable within a word (17a) and, in the case of a tie, stress falls on the rightmost non-final of the tied syllables (17b).

(17) Hindi stress (Kelkar 1968)

a.	ˈʃoːx.dʒa.baː.niː	'talkative'
	mu.sal.ˈmaːn	'Muslim'
	ru.pi.ˈaː	'rupee'
	ki.ˈd^{h}ar	'which way'
b.	ˈaːs.mãː.dʒaːh	'highly placed'
	kaː.ˈriː.ga.riː	'craftsmanship'

The Hindi weight hierarchy can be analyzed moraically (Hayes 1995; Broselow *et al.* 1997) if superheavy syllables are treated as trimoraic, heavy syllables as bimoraic, and light syllables as monomoraic. A constraint, Pk-Prom (Prince and Smolensky 1993; Walker 1996) selects the heaviest syllable in the hierarchy to be stressed, while a rightward alignment constraint working in conjunction with Nonfinality derives the correct results in case of a tie in weight.

In some languages, scalar weight is sensitive to vowel quality rather than syllable structure. For example, in Kobon (Davies 1980), stress falls on the heavier of the final two syllables, where weight is defined by the scale: Low Vowels > Mid Vowels > High Vowels > Centralized Vowels; in case of a tie, the data are less clear in showing which of the final two syllables is stressed. Kenstowicz (1997) proposes a series of prominence constraints referring to vowel quality to account for the Kobon facts. One constraint bans stress on low vowels, *P/a, another on mid vowels, *P/e,o, another on high vowels, *P/i,u, and another on central vowels, *P/ə,ɨ. The ranking of these constraints is fixed on a universal basis where constraints banning stress on less sonorous vowels are ranked above constraints banning stress on more sonorous vowels, that is, *P/ə,ɨ >> *P/i,u >> *P/e,o >> *P/a. Kenstowicz adopts a foot-based approach in which FtBin and a rightward alignment constraint ensure that a binary foot is constructed at the right edge of a word, where the type of foot, iambic or trochaic, is determined by the relative weight of the final two syllables.

More problematic for a coherent metrical stress theory are cases in which different parts of the weight hierarchy have different spheres of influence. For example, in Klamath (Barker 1964), stress falls on the penultimate syllable if it is either CVV(C) or CVC(C) (18a) and otherwise on the antepenultimate syllable (18b). However, a long vowel in final position or to the left of the antepenultimate syllable attracts stress away from CVC in both penultimate and antepenultimate position (18c):

(18) Klamath stress (Barker 1964)

a.	saˈɢapd͡ʒol	‘to play cat’s cradle’
	seˈsadwi	‘to sell’
	gaːˈmoːla	‘finishes grinding’
b.	ˈt͡ʃʼaw̯iga	‘is crazy’
c.	gaˈbaːtambli	‘goes back to shore’
	gaˈw̯iːnapbabli	‘is going among again’
	sakʼamsiˈneːʔ	‘to be lonesome’
	n̯isˈqʼaːk	‘little girl’

The stress facts suggest a three-way weight hierarchy with CVV being heaviest, since it can attract stress in any position, CVC being intermediate in weight, since it can attract stress in penultimate position, and CV being lightest, since it does not attract stress in either penultimate or final syllables. The complication posed by this system is that CVV can attract in contexts where CVC cannot, namely word-finally and to the left of the antepenultimate syllable. Thus, there is no single footing algorithm or alignment principle that will account for the distribution of stress involving both CVV and CVC.

Working within a rule-based paradigm, Hayes (1995) proposes that stress on CVV and CVC in the penultimate syllable and CV in the antepenultimate syllable reflects a right-aligned trochaic foot with the proviso that final syllables are not footed, that is, sa(ˈɢap)d͡ʒol, (ˈt͡ʃʼaw̯i)ga. A separate prominence tier treats CVV as heavier than both CVC and CV, thereby capturing the attraction of stress by final CVV and CVV to the left of the antepenultimate syllable. An undesirable feature of this analysis is that it assumes a separation between footing and prominence without offering a principled explanation for why two formally distinct devices should be invoked to account for the superficially uniform phenomenon of stress. On the other hand, Hayes (1995) notes that long vowels attract high tones in the intonation system, prompting him to suggest that the attraction of stress by long vowels may be tonally rather than stress-driven.

5 Metrical Stress Theory and Intonational Prominence

The interaction between stress and intonation in Klamath ties into a broader issue in metrical stress theory that has received relatively little attention in the theoretical literature compared to stress typology: the interaction between word-level stress and intonational prominence at higher prosodic levels (see Gussenhoven 2004, 2007 for an overview of this literature). While typological knowledge of the range of variation in stress systems appears to be relatively robust, it is unclear from most published descriptions of stress whether the described patterns reflect those found in words uttered in isolation or in a phrasal context. However, a long vowel in final position or to the left of a CVC penult or antepenult attracts stress

(18c). Fortunately, the body of literature discussing differences between phrase-level and word-level prominence is increasing.

One observation that appears to hold for most languages is that prominence is constructed bottom-up in such a way that phrase-level prominence patterns are overlaid on top of word-level stress (see Ladd 1996 for an overview). For example, in the English sentence *The elephants attacked the alligator*, the first syllable of *alligator* is the most prominently stressed syllable of the sentence when uttered under neutral declarative intonation. Contrastive focus on one of the other words can change the location of phrasal prominence. For example, if *elephants* is focused, as in the sentence *The elephants attacked the alligator, the birds didn't attack it*, the primary phrasal stress falls on the first syllable of *elephants*. If *attacked* is focused, as in *The elephants attacked the alligator, they didn't feed it*, the phrasal stress falls on the second syllable of *attacked*. Regardless of focus, phrasal stress selects the primary stress in the word for promotion (with the exception of cases of stress retraction to a secondary stressed syllable to avoid stress clash). Thus, one would not say *The elephants attacked the alligator* with phrasal stress on the second, third, or fourth syllables of *alligator* or the second or third syllables of *elephants* or the first syllable of *attacked*. Phrasal accent is typically phonetically manifested tonally in the form of a pitch accent, where the type of pitch accent, for example, H* or L* or a combination of the two, varies depending on the language and semantic properties of the utterance (Goldsmith 1978; Pierrehumbert 1980). Though an utterance may have multiple pitch accents, there is characteristically one, the "nuclear pitch accent," that is stronger than others.

In an Optimality-theoretic analysis, alignment constraints capture the attraction of the pitch accent by the primary stressed syllable of the phrase. One of these constraints captures the docking of the nuclear pitch accent on the rightmost content word in the default case. This constraint, formulated by Gussenhoven (2000) as ALIGN (T*, R, IP), requires the pitch accent to align with the right edge of an Intonation Phrase, a large prosodic constituent associated with a number of phonetic properties, including a final pitch excursion and final lengthening (Pierrehumbert 1980). Another constraint, ALIGN (T*, HEADσ_{IP}) (Gordon 2003), requires that the pitch accent falls on the main stressed syllable of the phrase.

Despite the strong cross-linguistic tendency for pitch accents to be associated with stressed syllables in bottom-up fashion, there is one phonetic property of utterance-final position that creates counter-examples to this tendency. Because phrase-final position is associated with a final pitch fall in the unmarked case, stress on phrase-final syllables is avoided in many languages (cf. Hyman 1977; Gordon 2000). This incompatibility between lowered pitch and stress creates mismatches in some languages, for example, Cayuga (Chafe 1977; Foster 1982; Michelson 1988), Onondaga (Chafe 1970, 1977; Michelson 1988), Seneca (Chafe 1977; Michelson 1988), Hill Mari (Ramstedt 1902), Central Alaskan Yupik (Leer 1985; Miyaoka 1985; Woodbury 1987), Tiberian Hebrew (Prince 1975; McCarthy 1979a; Dresher 1980; Rappaport 1984; Churchyard 1989, 1999), Chickasaw (Gordon 2003), between stress patterns of phrase-medial words where stress falls on final syllables, and prominence patterns in phrase-final words where stress is rejected

by final syllables. These cases potentially represent instances of top-down assignment of phrasal prominence, since word-level stress conventions that place stress on final syllables are overridden by considerations specific to phrase-final position. This positionally governed asymmetry can be formally modeled by assuming a context sensitive NONFINALITY constraint specific to phrase-final position in addition to a generic NONFINALITY constraint relevant for all words regardless of prosodic context. Alternatively, the intonational motivation behind final syllable avoidance at the phrase level can be modeled with direct reference to tones, adopting Pierrehumbert's (1980) autosegmental/metrical model of intonation in which phrase-final position is associated with a boundary tone, an idea originally proposed in Sag and Liberman (1975). Pursuing this approach, final stress avoidance in phrase-final position is captured, following Gussenhoven (2000) and Gordon (2003), by a constraint against crowding of two tones, the boundary tone and the pitch accent, onto a single syllable. If this constraint, *CROWD Gussenhoven (2000), is ranked above ALIGN (T*, R, IP), then the pitch accent, and stress along with it, retracts onto the penultimate. If the opposite ranking of *CROWD and ALIGN (T*, R, IP) obtains, then the pitch accent (and stress) is free to fall on the final syllable. This account assumes that the rightmost stress in a phrase receives a pitch accent.

The tonal account of final stress avoidance finds support from Chickasaw (Gordon 2003). In Chickasaw, the nuclear pitch accent falls on the final syllable of statements, which lack a final boundary tone, but on a pre-final syllable under most circumstances in questions, which are marked by a final low boundary tone. This split between utterance types differing in their boundary properties is exactly the type of asymmetry predicted by a metrical account based on tonal factors.

6 Foot-based vs. Grid-based Theories of Stress

An alternative to foot-based theories of stress is to model prominence as a rhythmic grid structure without any word-internal constituent structure larger than the syllable, an idea proposed by Liberman and Prince (1977) and developed by Prince (1983), Selkirk (1984b), and others. Under this approach, the grid consists of a sequence of strong, that is, stressed, and weak, that is, unstressed, syllables and distinctions in degree of stress, that is, primary vs. secondary stress, are captured in terms of differences in level of prominence. For example, the representation of the Maranungku word ˈjaŋarˌmata 'the Pleiades', in a grid-based approach to stress would be as in (19).

(19)

Level 1 (Primary stress)	x	.
Level 2 (Secondary stress)	x .	x .
	ˈja ŋar	ˌma ta

The lower level of stress reflects the secondary stress, where any syllable dominated by an "x" only on the lower grid level has secondary stress, for example,

the third syllable in (19). The higher grid level captures primary stress, which falls on the first syllable of the word.

Grid-based theories and foot-based accounts largely overlap in their empirical coverage of stress systems, though their predictions differ in some ways. As a departure point for our comparison of the two approaches, let us briefly consider Gordon's (2002a) grid-based Optimality-theoretic account of stress, which builds on earlier grid-based approaches couched within a derivational framework (Prince 1983; Selkirk 1984). As in the foot-based OT theory, leftward and rightward alignment constraints play an important role in accounting for the location of stress in all types of stress systems. In a grid-based theory, these two constraints, formalized here simply as ALIGN-L and ALIGN-R, refer directly to stress rather than to feet. Binary stress is attributed to two rhythmic constraints ensuring that stress falls on alternating syllables: *LAPSE bans adjacent unstressed syllables, while *CLASH prohibits a sequence of stressed syllables. Ternary stress is attributed to the high ranking of a constraint banning sequences of three unstressed syllables, *EXTENDED LAPSE, and the low ranking of *LAPSE. The directionality of the parse in binary and ternary stress systems depends on the relative ranking of the alignment constraints as in the foot-based theory.

In a grid-based theory, the relevant constraint needed to capture weight-sensitive stress like that observed in Chickasaw is the OT counterpart to Prince's (1990) Weight-to-Stress Principle, which requires that heavy syllables be stressed. The WEIGHT-TO-STRESS PRINCIPLE (WSP) is violated by a heavy syllable that is not stressed. WSP dominates *CLASH as well as the right alignment constraint that accounts for the normal unstressed-stressed alternating pattern found in words without heavy syllables. This is illustrated for a Chickasaw form in (20).

(20) Weight-sensitive stress in Chickasaw

tʃonkaʃ 'heart'	WSP	ALIGN-R	*CLASH
☞ ˌtʃon'kaʃ		*	*
tʃon'kaʃ	*!		

A difference between Gordon's grid-based theory and the foot-based analysis emerges in the treatment of peninitial, penultimate, and antepenultimate stress. In the foot-based account, these three patterns reflect a foot aligned either with a word edge in the case of penultimate and peninitial stress or one syllable removed from the right edge in the case of antepenultimate stress. Penultimate and antepenultimate stress are attributed to a trochaic foot and peninitial stress to an iambic foot. In Gordon's grid-based approach, penultimate, peninitial, and antepenultimate stress result from competition between alignment constraints and three anti-lapse constraints localized to the edge of words. *LAPSE RIGHT bans adjacent unstressed syllables at the right edge of a word, *EXTENDED LAPSE RIGHT prohibits three consecutive unstressed syllables at the right edge, and *LAPSE LEFT

militates against a sequence of two unstressed syllables at the left edge of a word. For example, ranking *Extended Lapse Right above Align-L yields the antepenultimate stress pattern of Macedonian, as illustrated in (21).

(21) Antepenultimate stress in Macedonian

vodenitʃari 'miller (pl.)'	*Ext Lapse Right	Align-L
☞ vode'nitʃari		**
'vodenitʃari	*!	

Similarly, ranking *Lapse Left above Align-R produces the peninitial stress system of Koryak, and ranking *Lapse Right above Align-L creates the penultimate stress pattern of Nahuatl.

6.1 Degenerate Foot Effects in Foot-based and Grid-based Theories

In a foot-based metrical theory, monosyllabic or "degenerate" feet, are predicted to arise only if there is a stray syllable remaining after all other syllables have been parsed into disyllabic feet. For example, in the five syllable Maranungku word ('ŋalti)(ˌriti)(ˌri) 'tongue', the final syllable forms a monosyllabic foot since there are no other unparsed syllables with which it can be grouped into a canonical disyllabic foot. Constructing a degenerate foot constitutes a violation of the requirement that feet be binary but has the virtue of ensuring that a foot occurs at a word edge, thereby honoring either Align-PrWd-Left or Align-PrWd-Right. However, because Align-PrWd-Left and Align-PrWd-Right refer to feet rather than directly to stress, they cannot ensure that stresses occur at a word edge. This means that stress systems that have stress on both the initial and final syllables are problematic for the foot-based theory. For example, the hammock stress systems with initial and final stress in Armenian (Vaux 1998), some dialects of Udihe (Kormushin 1998), and Canadian French (Gendron 1966) involve stress on both syllables in disyllabic words. This pattern necessitates a parse into two monosyllabic feet, that is, ('σ)('σ), which are always dispreferred to a single disyllabic foot. The initial plus final stress pattern is also problematic for the foot-based theory for another reason. In words of at least four syllables, a trochaic foot must be assumed at the left edge of the word and an iambic foot at the right edge; it is therefore unclear how this split type of foot system could be generated.

There are also at least two binary stress languages that consistently stress both the initial and final syllable across words of different lengths. In Tauya (MacDonald 1990), stress falls on odd-numbered syllables from the right plus the initial syllable. In foot-based terms, this system entails an iambic pattern from right to left plus two degenerate feet at the left edge of even parity words: ('σ)('σ)(σ'σ)(σ'σ). The sequence of two monosyllabic feet at the left edge of the

word parse cannot readily be generated by the proposed constraints, which predict that a parse into a single binary foot is always preferable to one into two degenerate feet. A similar difficulty arises for the foot-based approach in generating the Gosiute Shoshone (Miller 1996) pattern, which differs minimally from the Tauya pattern in employing a binary pattern starting from the left edge and displaying a clash at the right edge in even parity words: ('σσ)('σσ)('σ)('σ).

Because they refer directly to stress rather than feet, grid-based accounts have the potential to capture stress systems with both initial and final stress. Gordon (2002a) posits a constraint, ALIGN EDGES, that requires the initial and final syllable to be stressed. This constraint competes with others, including NONFINALITY (Prince and Smolensky 1993), but is unviolated in languages with both initial and final stress, including those with hammock patterns, that is, Canadian French, Armenian, and Udihe, and those with binary patterns, that is, Tauya and Gosiute Shoshone.

6.2 *Segmental Correlates of Metrical Structure*

Despite the grid-based account's success in deriving stress at word peripheries, the absence of feet in the grid-based theory may also be a shortcoming, as there is evidence that certain segmental processes are sensitive to the foot as a constituent. One such phenomenon, considered in Section 6.2.1, involves a well-documented asymmetry between the prosodic profile of trochaic feet and iambic feet that emerges in the examination of quantity-sensitive stress. Another piece of evidence for foot structure that is discussed in Section 6.2.2 comes from segmental alternations that are predictable from feet rather than stress.

6.2.1 The Iambic/Trochaic Law There is an interesting feature of iambic systems that differentiates them from trochaic ones. Many iambic stress languages lengthen stressed syllables in a foot, thereby creating feet consisting of a light syllable followed by a heavy syllable. Chickasaw is one such language, with its pattern of vowel lengthening in stressed non-final open syllables (Munro and Ulrich 1984; Munro and Willmond 1994). Thus, the second and fourth vowels in the word /tʃipisa**li**tok/ 'I looked at you' undergo lengthening to produce the surface form (tʃi,pi:)(sa,li:)('tok) 'I looked at you'. Patterns of rhythmic lengthening of stressed syllables are less prevalent in quantity-sensitive trochaic systems (Mellander 2001), particularly those applying to all stressed syllables in a word and not just the primary stress. This apparent difference between trochaic and iambic stress systems supports an asymmetric foot inventory in which iambs are inherently biased in favor of quantity-sensitivity (Prince 1990). In fact, certain trochaic stress languages display shortening of stressed syllables, which creates feet consisting of two syllables of roughly equivalent duration, that is, foot-internal isochrony. For example, in Fijian (Schütz 1985), phrases that underlyingly end in a heavy penultimate syllable (one containing a long vowel) followed by a light ultima shorten the long vowel in the penultimate; thus, underlying /m͡bu:ŋ͡gu/ 'my grandmother' surfaces as 'm͡buŋ͡gu (Schütz 1985: 528). Hayes (1995a), following

the spirit of a proposal advanced by Schütz (1985), offers an explanation for this a priori anomalous process of stressed syllable shortening in terms of the rhythmic principles underlying the trochaic foot. In Hayes's analysis, the penultimate and the final syllable initially form an unbalanced trochaic foot, which then shortens in order to create a trochaic foot ('m͡buŋ͡gu) consisting of two syllables of equivalent weight. By shortening the stressed syllable of the foot the pressure for durationally balanced trochaic feet is satisfied.

Interestingly, a bias toward durationally unbalanced light plus heavy iambs in contrast to durationally balanced light plus light trochees finds support from not only stress systems, but also psycholinguistic experiments, music and poetry (see Hayes 1995a: 79–81 for overview). In the non-foot-based theories of stress proposed thus far, any asymmetry between the durational patterns observed in trochaic stress systems and those found in iambic systems is an accident. The iambic/trochaic asymmetry may thus constitute one of the strongest pieces of evidence for the foot.

6.2.2 Mismatches Between Stress and Metrical Structure Another piece of evidence for the foot comes from segmental processes that are predictable from constituent structure rather than stress. The existence of segmental fortition in stressed syllables and lenition in unstressed contexts is well known. For example, Dutch (Booij 1995) inserts an intervocalic glottal stop as an onset to stressed vowels; epenthesis does not interrupt vowel sequences in which the second vowel is unstressed. We thus have pairs such as ˈxa.ɔs 'chaos' and a.ˈʔɔr.ta 'aorta' in which the presence of glottal stop is predictable based on stress. Similarly, English reduces most vowels to schwa in unstressed syllables.

Interestingly, there are also segmental alternations that are not predictable from stress but that fall out from analyses assuming metrical constituency distinct from stress. Vaysman (2009) documents several cases of this type, one of which we consider here. In Nganasan (Tereshchenko 1979; Helimski 1998; Vaysman 2009), a Uralic language, primary stress is confined to a three syllable window at the right edge of a word. Stress falls on a final long vowel or diphthong (22a), and on the penultimate if it contains a full, that is, non-central, vowel (22b). If the penultimate contains a central vowel and the antepenultimate contains a non-central vowel, stress falls on the antepenultimate (22c). Secondary stress falls on heavy (CVV) syllables and odd-numbered syllables counting from the left edge of a word except if this would entail a clash with another stress (22d).

(22) Nganasan stress (Vaysman 2009)

a.	ky'maː	'knife'
	le'hua	'board'
b.	ba'kunu	'salmon'
	tɨ'rimi	'caviar'
c.	ku'butənu	'skin, fur (prol., non-possessive)'
	'hʷaːtənu	'tree (prol., non-possessive)'

d.	ˌkita'raː	'only cup'
	ˌbakuˌnumə'numə	'my salmon (prol.)'
	ˌkaðarmə'numə	'my light (prol.)'
	suː'ðəːðu	'his/her/its lung'

Nganasan also displays an alternation between intervocalic consonants, termed "consonant gradation," whereby strong consonants, typically either voiceless or prenasalized obstruents, alternate with weak consonants, which are in most cases voiced or not prenasalized. The appearance of strong and weak pairs of consonants, indicated by boldface in the examples below, is predictable from syllable count (23). In the onset of even-numbered non-initial syllables, the strong grade appears, while the weak grade appears in the onset of odd-numbered syllables. Long vowels interrupt the alternating syllable count and, as long as they are not word-initial, are always preceded by weak consonants.

(23) Nganasan consonant gradation (Vaysman 2009)

ˌjama'ða-**t**u	'his/her/its animal'
ˌŋoru'mu-**t**u	'his/her/its copper'
ˌsuː'ðəː-**ð**u	'his/her/its lung'
ŋu'hu-**ð**u	'his/her/its mitten'

As Vaysman shows, this pattern is explained if one assumes that words are parsed into binary feet starting at the left edge of words with long vowels forming monosyllabic feet and degenerate feet allowed word-finally. Strong consonants occur foot-medially and weak consonants occur in foot-initial syllables that are not also word-initial (24).

(24) Nganasan consonant gradation and foot structure (Vaysman 2009)

(ˌjama)('ða-**t**u)	'his/her/its animal'
(ˌŋoru)('mu-**t**u)	'his/her/its copper'
(suː)('ðəː)-(**ð**u)	'his/her/its lung'
(ŋu'hu)-(**ð**u)	'his/her/its mitten'

The interesting feature of the Nganasan data is that stress does not always fall on syllables predicted to be stressed by the metrical structure diagnosed by consonant gradation. For example, the final foot in the last two words in (24) and the first foot in the third word have no stress, while the first foot in the last word is iambic rather than trochaic, unlike the other bisyllabic feet in the words illustrated here.

In order to account for this mismatch between metrical structure and stress in Nganasan, Vaysman adopts a hybrid approach employing both feet and grid-based prominence representations. In some cases, constraints referring to feet and those referring to grid marks conspire to produce alignment between grid marks and foot structure. For example, the location of secondary stress is largely consistent

with the foot structure diagnosed by consonant gradation. In other cases, however, grid marks fail to dock on syllables that are metrically prominent according to the foot structure. The reader is referred to Vaysman (2009) for the details of her analysis.

7 Metrical Stress Theory and Formal Acquisition Models

The constraint-based paradigm of Optimality Theory offers the opportunity for developing computational algorithms designed to model the language acquisition process. Some of these models have been applied to simulate the learning of stress systems (Tesar and Smolensky 2000; Tesar 2004, 2007; Heinz *et al.* 2005; Heinz 2006; Hayes and Wilson 2008). One such model is the maximum entropy model of phonotactic learning employed by Hayes and Wilson (2008). In their model, constraints are weighted in their relative strength rather than operating within a strict dominance hierarchy as in classical OT. The weightings of each constraint are constantly reassessed by comparing the number of expected violations of that constraint given the current weightings with the observed number of violations in a corpus of data. Hayes and Wilson also implement a model of constraint learning which discovers the constraints that account for the observed data. They submit their model of constraint learning and weighting to several empirical tests including the acquisition of two types of stress systems. Stress systems are of particular interest as a test of a learning algorithm since they represent a type of non-local phonotactics where positioning of stress requires examination of strings larger than adjacent syllables. First, they tackle the default-to-opposite sub-type of unbounded stress system, in which stress falls on the heavy syllable closest to one edge of words with at least one heavy syllable, but on the light syllable at the opposite edge in words lacking heavy syllables. For example, stress in Chuvash (Krueger 1961; but see Dobrovolsky 1999 for an alternative account) is reported to fall on the rightmost heavy syllable, that is, syllable containing a non-central vowel, otherwise on the initial syllable in words with only light syllables. Hayes and Wilson make the assumption that all heavy syllables (see Section 4.2) and all initial syllables in a language with this type of stress system are stressed and that End Rule Right promotes the rightmost stress in a word to primary stress. Employing a set of six predicates capturing the dimensions of syllable weight and stress (L, ˌL, 'L, H, ˌH, 'H), their training data consist of all combinations of symbols conforming to the stress description for words consisting of up to five syllables. Given this data, their grammar discovered five constraints, each with a weighting: one requiring every word to have a single primary stress, another requiring that the rightmost stress be the main one, another banning unstressed heavy syllables, another requiring the initial syllable to be stressed, and one banning stress on non-initial light syllables. Given these constraints and their weightings, various stress patterns can be assigned grammaticality scores that are calculated as a function of the number and weighting of constraint violations. In testing their

model, Hayes and Wilson find that all of the legal forms for their sample default-to-opposite stress pattern received perfect grammaticality scores, whereas other illegal forms received substantial penalties. In a separate test, Hayes and Wilson feed the 33 quantity-insensitive stress systems in Gordon's (2002a) typology to their learner, which discovered six constraints. They test the 33 learned stress grammars by submitting all possible strings of eight or fewer syllables to their model. In support of their algorithm, they find that all licit stress patterns received perfect scores whereas the illicit ones incurred substantial penalties. Similarly successful results obtained in the acquisition of the weight-sensitive stress system (as well as other phonotactic properties) of the Australian language, Wargamay.

A key feature of Hayes and Wilson's tests of the acquisition of metrical structure is their adoption of a grid-based rather than a foot-based approach to stress. This avoids the burden of having to model the learning of the hidden structure implicit in a foot-based theory. Whereas a language learner can hear stresses, an additional leap is required to infer the foot structure behind the stresses. Indeed, a single stress pattern can reflect multiple foot parses. For example, the Maranungku pattern with stress on odd-numbered syllables counting from the left could reflect the following foot structures in a four syllable word even if one constrains the number of possible parsings by assuming a ban on feet larger than two syllables: ('σσ)(ˌσσ), ('σ)(σˌσ)σ, ('σσ)(ˌσ)σ, ('σ)σ(ˌσ)σ. The learner must choose the correct parse that is also consistent with the footing in words with differing numbers of syllables. Though this is not an insurmountable difficulty for a learning model, it presents challenges, as it presumably would for a child as well (see Tesar and Smolensky 2000; Tesar 2004, 2007 for approaches to the acquisition of hidden structure).

8 Conclusions

The formal analysis of stress systems has provided fertile ground for phonologists working within both rule-based and constraint-based frameworks. Several issues remain unresolved, however, independent of the paradigm one assumes, including the role of the foot in stress systems and the interaction between word-level stress and intonational prominence. The advent of Optimality Theory has also raised issues of direct relevance to metrical stress theory, such as the formal modeling of the acquisition of stress.

ACKNOWLEDGMENTS

The author wishes to thank John Goldsmith, Jason Riggle, and Alan Yu for their numerous thoughtful comments on earlier drafts of this paper.

6 The Syllable

JOHN GOLDSMITH

1 Overview and Brief History

1.1 Introduction

In 1968, Ernst Pulgram began his classic monograph on the syllable with the wise words, "conscience, courtesy, and caution require that anyone wishing to concern himself with the syllable read all, or at least most, of the enormous literature on it." The years since his study have only magnified the challenge of this suggestion. Yet life is short, and space shorter still, and so in these few pages, I will attempt to survey the range of beliefs, models, and theories regarding the syllable that have been held by linguists, and attempt to integrate and compare them. Among the studies that I have found particularly useful are those by Fischer-Jørgensen (1952), Pulgram (1970), Fudge (1969), Goldsmith (1990), Blevins (1995), van der Hulst and Ritter (1999b), and Tifrit (2005), and my goal is only to supplement them, not to replace them. I have emphasized here the historical development of the approaches to syllable structure proposed over the past century insofar as it is relevant to today's phonologist, and I assume the reader has at least a basic familiarity with the role and usefulness of syllables in phonological analysis. For a description of a wide range of phenomena associated with syllabification, the reader is invited to consult Blevin's chapter on the syllable in the first edition of this Handbook, or any of the other references just cited.[1]

The syllable is one of the oldest constructs in the study of language, and most studies of phonology have found a place for the syllable within them. The momentous reconstruction of the behavior of Indo-European sonants, which was the greatest accomplishment of nineteenth-century linguistics, was intimately linked

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

to the realization that certain segments could be realized in strikingly different ways, depending on the location in which they appeared in their syllable: elements that could be identified as glides, nasals, and liquids would be realized as consonants in some contexts, but in others, when a vowel was not present for morphophonological reasons, the segment would be realized as a syllabic peak. Working out the solution to problems of historical linguistics led directly to the development of new conceptions of phonological structure, a historical event that has not yet been completed. This chapter is an overview of the evolution of the discipline's thought on this subject.

Tradition has it that a syllable consists of a vowel, usually preceded by one or more consonants, and sometimes followed by one or more consonants. In the overwhelming majority of spoken languages (though perhaps not all),[2] the syllable plays an important role in analyzing phonological regularities that phonologists have placed at the center of the phonological stage. The syllable is, first of all, important for the expression of statements of phonotactics, the principles of a language that describe which strings of basic sounds are found. Why does *blick* appeal to the anglophone more than *bnick*? It is, additionally, relevant for the expression of phonological conditioning for the realization of the basic sounds: the description of the realization of a *t* in American English is far more compact if the description can use the notions of "syllable onset" and "syllable coda" than if it is forced to forego them. Finally, there are few languages in the world whose prosodic systems can be adequately and compactly characterized without making reference to the syllable.[3] Prosodic or suprasegmental regularities involve a wide variety of linguistic phenomena, including timing, other rhythmic effects (such as clapping and dancing), and tonal structure.

As we shall see over the course of this chapter, the study of the syllable in recent decades has been an integral part of the development of theories of phonological representation, and to a lesser degree a part of the development of the theory of rules, constraints, and their interactions. But the study of the syllable has moved in fits and starts, with movement in several directions, all at the same time. While our knowledge, on the whole, has grown, it has done so by pursuing several different ideas.

We offer in this chapter a classification of approaches to the syllable; this is its only point of originality. There are several principal views of the syllable that have dominated linguistic discussions. Each has given rise to one or more formal models which encapsulate what is appealing about it. In the end, good reasons have been adduced for each approach, as I shall try to show, and it is hardly surprising, therefore, that much work has glossed over the differences between these models.

Of the general approaches, the first, that with the greatest longevity, is the view that sees spoken language as organizing sounds into wave-like groupings of increasing and then decreasing sonority (whatever *sonority* may turn out to be), while the second view sees the chain of segments of the language as organized into constituents in a fashion similar to the way in which words of a sentence are organized as constituents. The third view focuses on local conditions on sequence

segments, seeing the syllable as a term we use to summarize the recurring pattern of segment possibilities over the course of a word or an utterance. There are two additional views. Of these, the first is the view that the syllable, rather than the segment, is the right level of analysis for production and perhaps for perception, and linked to this is the hypothesis that while syllables are inherently ordered in time, the linear ordering of segments within a syllable may be the result of general principles of construction of syllables.[4] The second is the view that the *dynamics* of spoken language crucially depends on the syllable. This can be approached in more than one way, to be sure. Any account of speech production must offer some account of the length and timing of the sounds produced by a speaker, and framers of hypotheses have often been tempted to establish as a principle the notion that languages tend to preserve an *isochrony* – a common temporal interval – between syllables, or between stressed syllables. More recent work has developed accounts of the syllable based on models of the temporal coordination between consonantal gestures contained in the syllable onset and the gesture producing the syllable's vocalic nucleus; see, for example, Browman and Goldstein (1988), Fujimura (1992).

In this chapter, we will survey the fortunes of the these accounts, but emphasizing the first three: *sonority view*, the *constituency view*, and the *segment sequence view* of the syllable, and say just a bit about the issues of articulatory gestures and of rhythm; we give none of them the full exposure they deserve, and focus on the three that are more closely tied to questions of phonological representation. It is a remarkable fact that each of these views has flourished, grown, and developed over the last 50 years, and the extant literature has not made a great effort to give a synoptic perspective on the approaches that have been taken; we hope to fill this gap. In the final section, we return to this observation, and ask whether it is a Good Thing or a Bad Thing that largely incompatible perspectives have flourished. Should not the True View eliminate the two other views after a certain period of time? We try to offer an answer to this question in the final section.

1.2 Waves of Sonority: Whitney and Saussure

The oldest perspective on the syllable may well have been inspired by the observation that the jaw opens and closes as one speaks. This perspective, the *sonority* view of the syllable, is based on the view that each segment in an utterance has a sonority value, and that there are *crests* and *troughs*, or *peaks* and *valleys* of sonority in the speech chain, with peaks coinciding with vowels and syllable nuclei, and troughs coinciding with boundaries between syllables.

If sonority rises and falls in the course of an utterance, we might expect to find a difference in the realization of consonants depending on where in the wave of sonority they appear. On this view, a consonant that appears in a context of *rising* sonority at the beginning of a syllable – that is, before the peak of the syllable – is in a qualitatively different environment compared to those that appear in the context of *falling* sonority, at the end of a syllable. Following terminology that goes back more than fifty years, we will call the first context, that of rising sonority,

the *onset*, and the second, that of falling sonority, the *coda*, and the peak of sonority the *nucleus*.

This view was well developed by the end of the nineteenth century, and is described in considerable detail in Whitney and Saussure. For example, Whitney wrote in 1874:

> The ordinary definition of a syllable . . . amounts to this: a syllable is that part of a word which is uttered by a single effort or impulse of the voice. Such an account of the matter is of only the smallest value . . . The governing principle, it seems plain to me, which determines [syllabification], is that same antithesis of opener and closer sounds upon which the distinction of vowel and consonant is founded. The vowel sounds of *any* are practically identical with those that compose our ē (the "long a" of *they* . . .); and ē may be so protracted so as to occupy the whole time of *any*, without giving the impression of more than a single syllable; but put between the two opener vowel elements the closer consonantal *n*, and the effect is to divide them into two parts: the ear apprehends the series of utterances as a double impulse of sound. So in *lap* there are three articulated elements, of three different degrees of closeness, but the *a* (*æ*) is so much more open than either of the others that they are felt only as its introductory and closing appendages; there is a crescendo-diminuendo effect, but no violation of unity. And *alp* and *pla*, in like manner, are a *crescendo* and *dimenuendo* respectively . . . [W]hen it comes to allotting to the one or the other syllable the closer sounds which intervene between the opener, there is room for much difference of opinion . . . Thus, for example, in *any*, the intervention of the *n* between the two vowels makes the dissyllable; but the *n* itself belongs as much to the one syllable as the other . . . There is, on the other hand, more reason for assigning the *p* of *apple* (*æ-pl*) to the second syllable . . . There are . . . sounds so open that they are always vowels, never occupying the position of adjuncts in the same syllable to another sound which is apprehended as the vowel of the syllable. Such is especially *a*; and *e* and *o* are of the same character. But *i* and *u* . . . become *y* and *w* on being abbreviated and slighted in utterance . . . Vowel and consonant are the two poles of a compound series, in which are included all the articulate sounds ordinarily employed by human beings for the purposes of speech. (pp. 291ff.)

Saussure[5] was centrally concerned with the realization of the three phonological classes of segments in Indo-European: those that were always realized as consonants (what we would today call the *obstruents*), those that were always realized as vowels (*non-high vowels*), and those which were realized in the one way or the other, depending on their context (*sonorants*); this third group consisted of liquids, nasals, and glides/high vowels. Saussure (1995: 222) established a set of four ordered rules for this, which aim primarily to account for the realization of sonorants, whose surface form is heavily dependent on the phonological context:

- Rule 1: a vowel that follows a sonorant puts that sonorant into an onset (Saussure's "explosive") position.
- Rule 2: an obstruent inhibits a sonorant that precedes it, as does silence. Here, something "inhibited" is syllabic, in modern terms.
- Rule 3: a glide which has become inhibitive has the same effect as a vowel.
- Rule 4: a sonorant which is in an onset has the same effect as an obstruent.

Saussure indicated explicitly that these rules must be applied sequentially from right to left, from the end of the word, and illustrated the effects on forms. The influence of his familiarity with the scholarship of Panini is evident.

As we will see, most theories of the syllable formulated with any desire for rigor have begun with a set of three or four basic syllabification rules, and the cross-theoretical appearance of such principles allows for a typology of sorts of approaches to understanding syllabification. We will see this below, in the work of Pulgram, Hooper, and Kahn.

The sonority hierarchy is discussed in some detail by phonologists early in the twentieth century, such as Jespersen (1904), van Ginneken (1907), and Jones (1918).

Fischer-Jørgensen expressed the prevailing perspective in her classic paper (1952), viewing the syllable "as a unit of speech containing one relative peak of prominence. The division of the chain of speech into syllables may be due simply to the inherent loudness of the successive sounds, but the peaks may be reinforced or altered by arbitrary changes of loudness, and this means may also be used to give a clear delimination of the units."

1.3 Constituents and Structure

1.3.1 Pike, Hockett, Fudge: The Arboreal View Immediately after World War II, two studies were published which proposed a new account of the syllable on the basis of the notion of *constituent*, a concept that was being developed at the frontier of syntax at the time: Pike and Pike (1947), and Kuryłowicz (1948).[6] This was a moment during which Bloomfield's notion of *constituent* was coming into general use in syntax, and to some extent replacing the earlier view of syntax, according to which syntax focused on asymmetric relations between pairs of words. On this newer constituent-based view, sentences were successively cut into smaller and smaller pieces, until reaching the word. Pike and Pike argued, why stop there? Why not continue to chop up the utterance into finer-grained pieces, since we already have a name for them: syllables! The momentous step of bringing insights from the domain of syntax into the treatment of syllable-level phenomena has continued to play a major role in the development of theoretical views, notably in approaches employing the concept of government.

Within American phonology during the post-War years, there were mixed feelings about the importance of the syllable for phonology, based to some extent on the fact that in the overwhelming majority of cases, syllabification is phonologically *predictable*: given a sequence of segments in a word in a specific language, the location of the syllable boundaries is predictable, which is to say, not distinctive. In a framework which required fully predictable information to be absent from the phonological representation such as was dominant in the United States at the time, it was reasonable to draw the conclusion that syllable structure should not be present in the phonological representation. On the other hand, syllable structure is probably the single most important conditioning environment for segmental rules (we will see examples of this in Section 2 below), so it is essential

for a theory to allow for the existence of a phonological representation in which syllable structure interacts with "choice," so to speak, of the phonetic spelling-out of underlying segments.

The question of the predictability of syllable structure is not as simple as suggested in the preceding paragraph, however. In some languages, syllabification appears to operate without any reference to morphological or word boundaries – as in Spanish – while in German, word boundaries and at least *some* morphological boundaries are critical to syllable establishment, and much the same is true of English, as we will see below (Section 2.3).[7]

1.3.2 Syntagmatic and Paradigmatic Syntacticians and phonologists agree on another point, one which is rather more abstract. It is this: when we study linguistic units – words in syntax, and segments in phonology – it is important to both analyze the specific relationship that a given unit bears to its neighbors in an utterance (and we call those syntagmatic relationships: subject, direct object; nucleus, onset), and the categories into which the inventory of units can be usefully subdivided (nouns, verbs, adjectives vocoids, fricatives, nasals). From the very earliest work we have cited, phonologists have borne in mind the fact that a particular item, such as a vocoid – a segment in which spectral resonances are its most salient properties – can be either the nucleus of a syllable, or not. In the former case, we call it a *vowel*, and in the latter we call it a *glide*. But using the convenient terms *vowel* and *glide* should not lead us to overlook where the constancy is and where the difference is between a vowel and a consonant.

Thus "being-a-vowel" – as opposed to "being-a-vocoid" – is a fact about the *role* a segment plays in a particular spot in an utterance: it is syntagmatic. How should this view, nearly universally held, be integrated into a larger or theoretical point of view? One negative conclusion is that it should not be as a binary feature (Selkirk 1984a makes this point explicitly, and she is not the only one to do so). What *is* it about its role in the syllable that is crucial: is it the fact that a syllable must have exactly one element that is nuclear or syllabic? Is this property something that inheres in the formal relationship between the syllabic element and the syllable constituent? Some have opted for this second alternative, and called the relationship *head*.[8] In constituency-based models, it is often assumed that among the nodes depending from a given node, there is exactly one node marked as its head, and the syllable nucleus is the unique element of the syllable that is a member only of constituents marked as head, within the syllable. In other models, the nucleus is identified as the most sonorous element. What is generally agreed is that the notion of *feature* is not best suited for this job: features, as an inherent non-relational object, are ill-suited for representing the important characteristics of differences that are inherently relational, and the difference between a vowel and a corresponding glide lies in the *relation* that exists between the segment and the context in which it appears.

1.3.3 How to Parse CVC The constituent model of syllabification naturally suggests that the syllable nucleus forms a constituent with either the onset or the

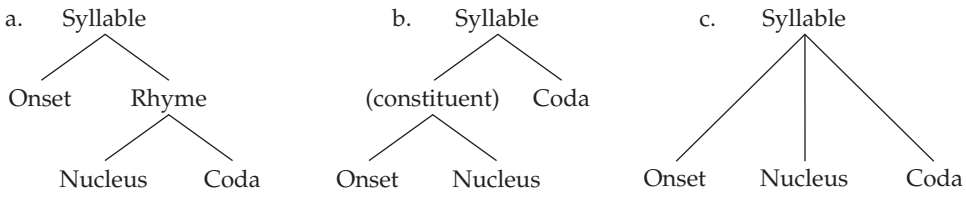


Figure 6.1 How to parse CVC.

coda (as in Figure 6.1): I say "suggests," because the flat structure of (c) is a possible analysis in a model with constituents.

Structure (a) in Figure 6.1 is the one most widely defended and used. The structure in (b) has been proposed on occasion. In Japanese, for example, Kuzobono has argued for structure of this sort (see Kubozono 1989a), for reasons that seem to be strongly linked to the centrality of the mora to Japanese phonological structure; Yi (1999) argues for a similar analysis of Korean, as did Bach and Wheeler (1981). There is a view that what has been treated as a CVC syllable should rather be analyzed as a CV syllable, followed by some kind of defective syllable, and such views involve strutures more like (b) than like (a) or (c) – a syllable with an onset but no nucleus; we discuss this below (Section 1.12). The structure in (c) has been proposed (by Saporta and Contreras 1962, cited by Harris, and defended in detail by Davis 1985); so-called "flat models" of syllable structure also come close to this; see the brief discussion below in Section 1.8.

Part of the difficulty in establishing a structure conclusively lies in the fact that there are few generally accepted principles for determining constituency in phonology, and those that do exist tend not to give decisive answers when applied to this question.

Some researchers have explored the relevance of language games and tasks that can be studied by means of psychological tests; in a series of papers (Stemberger and Treiman 1986; Treiman 1986; Treiman *et al.* 2000) Treiman and colleagues have taken this approach, and been influential in turn (see, for example, Stenneken *et al.* 2005 and references there). Some of this work is based on the suggestion that manipulations of linguistic segments require what we might think of as cutting and repasting pieces of syllables to form new syllables: for example, if /krɪnt/ and /glʊpθ/ are to be merged, will a speaker give us /krʊpθ/ or /krɪpθ/, or something else? If we formulate a binary-branching tree structure over the set of phones, then a location between any two adjacent segments can be associated with a height in the binary tree, and we may hypothesize that locations corresponding to "high" nodes are preferred as positions for cutting: the position between the last segment of the onset and the first of the rhyme corresponds in this sense to a higher position in the syllable than the position between the first and second consonants of a syllable onset. But hierarchical structure is not the only, and certainly not the most direct, model capable of making predictions as to preferred cut-points in psychological tasks; the fact that a piece of beef may be

easier to cut at some points than at others should not be taken as evidence that the meat is hierarchically organized.

Perhaps the most widespread principle which phonologists have attempted to apply is one that says that when adjacent segments are part of smaller constituents, there should be stronger cooccurrence restrictions bearing on the two segments than when the segments are more distantly related, that is, are parts only of larger constituents. On this account, if there is a rhyme constituent, as in [$_{Syllable}$*Onset*[$_{Rhyme}$*Nucleus Coda*]], then we should find stronger cooccurrence restrictions between the nucleus and the following coda segment than between the vowel and the (final) consonant of the onset (since the latter pair of segments are in a larger constituent, the syllable, but the former are in a smaller constituent, the rhyme). This perspective is clearly presented in Pike (1967), Fudge (1969), and Selkirk (1982), and appears to be originally rooted in the early reflections on the meaning of constituency, notably Rulon Wells's 1947 paper (Wells 1947); again, the principles established on the basis of exploring syntactic structure were applied, in retrospect with relatively little reflection, to problems of phonological representation. Davis (1985) discusses a number of weaknesses of this argument.

Unfortunately, as it stands, this is a rather vague formulation, and even if the notion of "strong cooccurrence restrictions" can be clarified (as it certainly can), it is not clear why constituent structure should cause the cooccurrence restrictions in question. This imprecise notion can be interpreted as standing in for the more precise and explicit measure of *mutual information*.[9] Mutual information measures the degree to which the frequency of any pair of successive segments departs from independence, that is, the degree to which the probability of a given pair of segments departs from the product of their independent frequencies. We can use such measures to determine whether pairs of segments that are structurally closer have, on average, greater mutual information, all this as a quantitative measure of the validity of the claim that syntagmatic structure in the syllable has an impact on the possibility of cooccurrence restrictions.

Harris (1983) argues that Spanish has a restriction limiting the number of segments in the rhyme to a maximum of three. He proposes this in a model in which the nucleus contains two segments in the case of a nucleus (e.g. *muerto*), and in other forms, the coda can contain two segments; and thus there is no purely local way of formulating the restriction that the coda has an upper limit of three elements. Harris is at pains to emphasize that he is aware of the mismatch between his finding and the ability of the hierarchical model to easily incorporate it (the problem comes from the apparent need to say that the rhyme cannot have more than one element in both the nucleus and coda, though it can have more than one element in either one separately).

1.4 Syllable Timing

Kenneth Pike (1947) was also responsible for the introducing the terms *stress-timed* and *syllable-timed* as descriptors of a language's rhythm. Pike suggested that some languages (he cited English) were stress-timed, and some were syllable-timed (he

mentioned Spanish). By this he meant that there was a strong tendency in a stress-timed language for stresses to appear equally spaced in time, while in syllable-timed languages, the tendency was greater for syllables to be equally timed (or *isochronous*). While this difference has survived many decades of usage, and a clear formulation in Abercrombie (1967: 97), it has not found experimental support over the years. Roach (1982) gives a brief summary of the issues raised and the difficulties encountered in dealing with the claims behind this distinction.

1.5 Classical Generative Phonology

During the classical period of generative phonology, many phonologists accepted the proposals in *The Sound Pattern of English* (1968), and SPE did not include the syllable within its linguistic *Weltanschauung*. On that all are agreed, but beyond it, there is little consensus regarding the relationship of the study of the syllable to generative phonology, and it is probably fair to say that there are no historical truths – only points of view. One widely held view is found, for example, in Féry and van de Vijver (2003a: 3): "In the seventies, several phonologists, such as Vennemann (1972a), Hooper (1976), and Kahn (1976), proposed including the syllable as a prosodic unit in generative phonological theory. The relevance of the syllable for linguistic theory has increased ever since." While this is certainly true, in the sense that it contains no false statements, it does suggest something that is not true, that the syllable has no relevant history before SPE. Peter Auer (1994) suggests that "credit for the restoration of [the syllable] is . . . due to a group of phonologists who in the 1970s, in schools such as Natural Generative Phonology, Natural Phonology, and Syllable Phonology, attacked orthodox MIT generative phonology (represented by *The Sound Pattern of English*) and whose inadequacies they showed."

1.6 Pulgram on the Syllable

Pulgram (1970), written in 1968 and published two years later, has had a great impact on the field despite the fact that it is relatively rarely cited explicitly. It offered a number of proposals (such as onset maximization) that are still widely adopted today (see Bell 1976 for an insightful review).

Pulgram offers a modern interpretation as to how to view the relationship between language-particular and universal characteristics of syllables and syllabification:

> If the syllable is an operative unit of all languages, it is also a universal of language. Its definition must be . . . the same for all languages, regardless of the varying unit inventories in the different [languages] . . . there arises the interesting question whether it might not be possible to arrive at a phonotactic definition of the syllable which . . . does have universal validity for all languages. The question is, in other words, whether the phonotactic rules on syllabation might not be formulated in such a way that they are applicable to all languages, even though their implementations

> in the different languages must differ because of the underlying differences of phonotactics. I believe that such general phonotactic rules on syllabicity are not only possible but also necessary for the proper syllabation of any utterance in any language. (p. 23)

His account is the first to offer a sequence of ordered rules for syllabification (p. 70ff.), beyond what we have seen Saussure had already proposed. After determining the size of the domain in which syllabification will be established (his "Rule 1"), his Rule 2 of *maximal open syllabicity* places a syllable boundary after each vowel. Rule 3, of *minimal coda*, says that if a phonotactic condition forbids certain vowels to appear syllable-finally, then the syllable boundary is shifted to the right (which is to say, one or more consonants are shifted "to the left" of the syllable boundary), but only the smallest number of shifts necessary to achieve a sequence that is possible word-finally. A sequence $C_1V_1C_2C_3V_2$ will be syllabified as C_1V_1–$C_2C_3V_2$ by Rule 2, and then as $C_1V_1C_2$–C_3V_2 if and only if V_1 may *not* appear syllable finally in the language and C_2 *is* a permitted syllable coda in the language. Rule 4, *maximal onset*, shifts the syllable boundary "to the right" (in the sense just discussed) if the syllabification so far has resulted in a syllable-initial sequence which cannot appear word-initially. Thus *e–mploy* is resyllabified by Rule 4 to *em–ploy*. Pulgram's Rule 5 (*principle of the irregular coda*) says that in case an interlude cannot be parsed into a legitimate word-final and a word-initial sequence, the coda must accept the material which it would not otherwise (i.e. word-finally) be forced to accept: "If the necessary transfer from syllable-initial to syllable-final position leads to an inadmissible syllable-final group of consonants, then the burden of irregularity must be borne by the coda rather than the following onset." (p. 51). Pulgram gives the example of Spanish *transcribir*, which he syllabifies as *trans–cri–bir*, despite the impossibility of word-final *ns* in Spanish.[10] Although, as we will see in Section 1.17, Fischer-Jørgensen had already documented a range of languages in which the set of interludes is significantly broader than the sequences of word-final plus word-initial sequences would suggest, Pulgram's core system does not allow for that, and he attempts (in this writer's view, unsuccessfully) to come to grips with that wrong prediction.

Pulgram's account is geometrically flat; the information that is contained in the correct syllabification of an utterance is, once the nucleus has been identified, nothing but a statement of where the boundaries are between syllables. One exception to that statement must be made, however: Pulgram emphasizes the importance of finding a notational means of expressing the idea that in English, a consonant may straddle a syllable boundary, which is to say, may be ambisyllabic.

1.7 Natural Phonologies

The syllable played a central role in natural generative phonology, described by Vennemann, Bybee, among others, and natural phonology, developed by Stampe and others (see Stampe (1972), and that played an essential role in the recrudescence

of work on the syllable. Hooper (1976) presents a range of arguments within an essentially formal and traditional generative perspective in favor of syllable analysis, utilizing a formal symbol (in the event, $, following Vennemann) among the string of phonological segments to mark syllable division. She notes, for example, that assimilation of nasals for point of articulation is found not simply when a nasal is followed by a segment specified for a point of articulation – but rather when the nasal is at the end of a syllable, and followed, in the next syllable, by an element with a point of articulation. She offers this as an explanation for the non-assimilation in words like *muevo* [mweβo] 'I move' despite the assimilation found in forms such as *un huevo* [uŋweβo] 'an egg'.

Hooper proposes an ordered set of rules for syllabification (p. 527): (1) Place a '$' between adjacent syllabics; (2) in a VCV sequence, insert a $ to form V$CV; (3) place a '$' before a sequence of obstruent followed by non-nasal sonorant plus vowel – though exceptions blocking this in the case of *tl, dl* are discussed, and a universal set of possible blocking conditions are proposed. Rule (2) is considered universal, while rules (1) and (3) are not – the corresponding rules in languages differ in specific ways that need to be made explicitly. Syllable theories deploying a boundary symbol to demarcate syllable boundaries can be understood either as claiming that's all there is, as far as syllables are concerned, and in particular there is no hierarchical structure, or they can be viewed as making modest positive claims, leaving open the possibility that there is further structure that the boundary symbol notation fails to indicate. Authors do not always make it clear which perspective they adopt. Selkirk (1982) notes a number of authors who use this notation (p. 354), but at least one of them (Hockett 1955) clearly indicated the presence of additional hierarchical structure.

1.8 Flat Structure

Kahn (1976) provided a number of convincing arguments for integrating the syllable into formal, generative accounts of phonology, and a number of influential studies within this framework followed, including, notably, Kiparksy (1979), Harris (1983), Selkirk (1982), Clements and Keyser (1983), all of which reflected in various ways both the formal concerns of classical generative phonology and the emphasis placed on the analysis of the syllable by natural and natural generative phonology. Kahn's work showed the usefulness of a formal model in which syllables were represented as symbols on a distinct tier, formally parallel to an autosegmental tier, an approach that made ambisyllabicity a natural notion. We will return to this in Section 2.3 below, when we consider the process of flapping in American English.

1.9 Metrical Phonology

Metrical phonology, which began as a theory of linguistic rhythm at the syllable level and above (Liberman 1975; Liberman and Prince 1977), quickly extended its domain to syllable internal structure (Kiparsky 1979; McCarthy 1979b), and

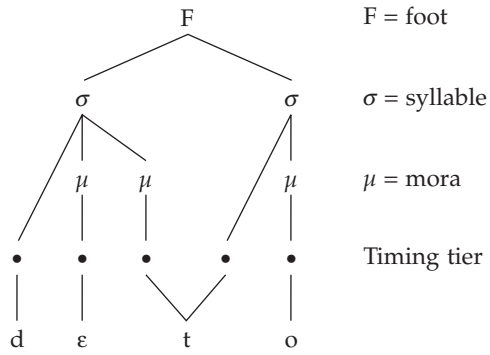


Figure 6.2 Timing tier.

provided the means to explore the possibility of hierarchical constituent structure within the syllable, extending the analysis that had begun with Pike, Hockett, Fudge, and others. This work was part of an intensive period of work on the theory of phonological representations, which also included work on autosegmental accounts of tone, quantity, and harmony. Most important for the theory of the syllable was the usefulness of autosegmental representation in understanding the nature of long (i.e. geminate) consonants and long vowels. This new perspective allowed an account with far fewer paradoxes: long segments would henceforth be analyzed as complex 2-to-1 representations linking a single segment (a consonant or a vowel) on one tier, with two segments on a tier whose elements represent temporal or rhythmic information, often called the *timing tier* for this reason. See Figure 6.2, where a long, or geminate, consonant is represented with a one-to-many association to elements on a timing tier, here represented as filled dots.

1.10 Sonority Redux

Hankamer and Aissen (1974) was an important reminder to the American phonological community that a gradient perspective on the distinction between vowels and consonants was often critical for understanding a phonological process, a message that the natural phonologists had also tried to communicate; see Stampe (1972), Hooper (1976), and Vennemann (1972). The literature at this time shows relatively little awareness that sonority analysis has deep roots in the linguistic literature, as we have seen, although some noted the role that sonority plays in Panini's grammar of Sanskrit. James Harris (1983: 15) refers to "the familiar sonority scale V > G > L > N > O", and (p. 21) employs this scale to establish the generalization that consecutive consonants in the Spanish onset must be of increasing sonority, and not adjacent on this sonority scale. Selkirk (1984a) argues for a replacement of the major class features (sonorant, consonantal, syllabic) by a variable, called sonority, that takes on values (perhaps limited to integers, perhaps not), and perhaps in the range [0,10].

A few years later, Dell and Elmedlaoui (see their 1985 and 1988) published a highly influential analysis of Imdlawn Tashlhiyt Berber (ITB) which presented the strongest argument to date for the importance of sonority in the treatment of syllabification, based on an elaborate set of principles which came down essentially to this. Syllabification is established by a sequence of ordered phonological rules, divided up into a set of "core syllabification" rules, followed by a further set of attachment rules. The core syllabification rules are a sequence of virtually identical rules that differ only in the sonority of the segment that they apply to; the rules all take the form: *Associate a [core] syllable with any sequence (Y)Z, where Z is a segment of type T*; and the algorithm passes through eight instantiations of that rule template, as the variable T proceeds through the segment classes: {the vowel *a*, high vocoids, liquids, nasals, voiced fricatives, voiceless fricatives, voiced stops, voiceless stops}, which is to say, as the variable T passes through the segment inventory from most sonorous to least sonorous.

The analysis takes into consideration the fact that in ITB, segments at any point in the sonority hierarchy can be nuclei of syllables, including even voiceless stops, in the right phonological context, and for any segment type, there are examples of segments in particular morphemes which alternate between being syllablic and being non-syllabic, based on the larger phonological context. Examples are given in Figures 6.3 and 6.4, where all and only the syllable nuclei are indicated with upper case letters (extending a notation employed by Dell and Elmedlaoui). As Dell and Elmedlaoui point out, the facts in ITB suggest that the traditional search for the language-specific boundary between possible and impossible syllables is misguided, or at least inadequate; it appears to be necessary to provide a notion of *preferred* syllable structure, since many of the incorrect syllabifications of words in ITB are composed of syllables, each of which would be possible if it occurred in some other context.

Proposing an analysis within a generative derivational account, Dell and Elmedlaoui encode their notion of preference by ordering a rule which assigns a preferred syllabification (e.g. segment plus low vowel) *earlier* in rule-ordering than a rule which assigns a less-preferred syllabification (e.g. segment + liquid), along

3m.sg.	*3f.sg.*	*Gloss*
IldI	tLdI	'pull'
IrbA	tRbA	'carry on one's back'
IxsI	tXsI	'go out (fire)'

Figure 6.3 IT Berber.

2sg.perfective	*3f.sg.perfective with dat. 3m.sg.object*	*Gloss*
tRgLt	tRglAs	'lock'
tSkRt	tSkrAs	'do'
tMsXt	tMsxAs	'transform'

Figure 6.4 IT Berber.

with a convention on rule application that blocks a later rule from syllabifying a segment that has "already" been syllabified. Other issues that seem to involve well-formedness of representations also enter into the formalism of rule application: for example, the *observation* that the only syllables without an onset are those that are word-initial becomes transformed into a constraint on syllable well-formedness; one that forces a rule not to apply if the rule's output would violate the constraint (e.g. underlying *iħaultn* 'he made them (m.) plentiful' becomes first *i* [*ħa*] *ultn*, as the sequence *ħa* forms a syllable; but then *u* does not become the nucleus of an onset-less syllable, because the result of that operation would violate the constraint on word-internal onsetless syllables, and the final syllabification is in fact [*I*] [*ħA*][*w L*] [*t N*]). Dell and Elmedlaoui also note that core syllabification appears to apply from left to right, in the sense that the correct syllabification of underlying *rksx* is [*R*][*kSx*] rather than [*Rk*][*sX*] 'I hid', but they note that the data is far from unambiguous on this point.

The analysis proposed by Dell and Elmedlaoui was thus a demonstration that it was not just the *inherent sonority* of the segments in a language, specified along an articulated scale, that could be the primary determinant of how syllabification was determined in a language, but the *relative sonority* of each pair of neighboring segments could play a crucial rule. Neither the constituency view of the syllable, nor the sequence-based view that we consider below, is capable of providing explicit formal means for making syllabification dependent on the relative sonority of segments in a string.

A number of generalizations that use the notion of sonority have been proposed to characterize common properties of syllables. Of these, the most important are the following:

- The Sonority Sequencing Generalization: sonority rises during the onset, and falls over the rhyme. Selkirk (1984a: 116), citing Hooper (1976) and Kiparsky (1979, 1981a), writes:

  > In any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values. . . . The SSG can be viewed as imposing universal constraints on the possible form of language-particular sets of conditions on syllable structure. It in no way constitutes on its own a theory of syllable phonotactics, however, for languages will differ precisely in their choice among the various conditions on terminal positions that are consistent with [the SSG].

- Minimum Sonority Difference (or dissimilarity). Steriade (1982: 94) proposes that once an appropriate numerical sonority hierarchy has been established, along the lines of Selkirk (1984a), a language may impose the restriction that adjacent segments must be a minimum sonority distance from each other.
- Dispersion Principle, proposed by Clements (1990): all other things being equal, a language will preferentially maximize sonority difference in the syllable onset, but minimize it in the coda.

But with all this in mind, what *is* sonority? The set of answers that phonologists have provided to this question range across a wide span of opinions as to what the ultimate object of phonology's study is. Some would suggest that its fundamental motivation lies in its association with physical energy, or with the degree of opening of the mouth, or both, while another phonologist, less concrete and more abstract, might offer a different answer: sonority is the name we give to our method of organizing the segments from a language along a one-dimensional scale, with the ultimate purpose of describing permissible syllables. This latter answer raises two questions immediately: if we could identify sounds independent of the language in which they appear, would it be the case that for any such pair s_1 and s_2 which can be found in several languages, their ordering would always be the same across languages? – that is, if *l* is more sonorous than *n* in one language, is the same true in every language that has both? A second, independent question is this: if it is indeed useful to compare the sonority of two segments s_1 and s_2 in a language by modeling with arithmetic values, would we want to say that s_1 has the same sonority in every environment, or could sonority be dependent on phonological context? We will return to this question in Section 1.13.

1.11 Worrying About Slots That Hang From Trees

The Pikean, arboreal view of syllables can be pushed to the point where the terminal nodes of the tree are viewed as playing a more important role in the theory than the segments do – if by "terminal nodes" we mean (as syntacticians often do) not the symbols referring to the phonological segments directly, but some sort of node that may be "empty" of any given segment; such a view leaves open the possibility of phonologically null elements that play a significant role in the model. In this section, we will review how this reversal has an impact on the treatment of onset clusters; in the next, how it leads to more radical statements about the nature of codas and empty nuclei.

In English and many other languages, an onset can consist of a single obstruent (*pa*) *or* sonorant (*la*), or it can consist of an obstruent plus a sonorant, in that order (*pla*). In order to account for two-segment clusters, the arboreal view instructs us to include a phrase-structure rule of the form *onset* $\rightarrow C_1C_2$, but what should C_1 and C_2 be? It seems easy to decide that C_2 should specify that the segment in this position should be a sonorant, but what about C_1? The problem is that phrase structure rules are well equipped to deal with generating sets of strings like {Ø, *p*, *b*, *l*, *r*, *pr*, *bl*, *pl*, *bl*}, but phonologists are not always satisfied with the result that they produce. A simple phrase-structure analysis is given by the following rule:

(1) $Onset \rightarrow \left(\left\{\begin{matrix} p \\ b \end{matrix}\right\}\right)\left(\left\{\begin{matrix} r \\ l \end{matrix}\right\}\right)$

But the phonologist may recoil at the conclusion that seems to follow from this: that the position of the onset consonant is different in *pa* and in *la*, that there is

no single statement that says that every class of segments can appear in the onset, and that the ordering of the obstruent and sonority must be specified in this language-particular fashion even though it reflects a widespread property of languages (e.g. Selkirk 1982: 346). Ultimately it is not clear how weighty any of these three considerations are. For example, we might argue that the first is not a valid conclusion to draw. Rule (1) will generate $[p]_{onset}$ and $[r]_{onset}$, and not something with an non-terminal node distinguishing between the two – that is, there is nothing like this to worry about: $[\ [_{obstr}\ p]\ ]_{onset}$. To put the same point more informally, phrase-structure rules do not create slots; they determine the set of admissible tree labelings.

1.12 Government Relations

Beginning in the 1980s, a number of phonologists have explored the consequences of encouraging syllable representations with empty nuclei, along with strict constraints on what can appear in onset or coda position – constraints which have the effect not just of *allowing* for the possibility of empty nuclei, but of strongly *requiring* a wide use of empty nuclei in the analysis of real data. For example, the constraint that there be no codas at all (Harris and Gussman 1998), and the constraint that there may be only one consonant in an onset, has the consequence that there are at least as many syllables as there are consonants in a given utterance. Proponents of this view criticize earlier perspectives as being too bound to representations whose terminal elements are the observed phones of the utterance:

> . . . we should first raise a fundamental question about the central premise of the phoneme-centred view: is it really the case that syllable structure is projected parasitically from segment strings? Suppose we entertain the alternative idea that syllable structure should be defined independently of segment strings and word structure. What empirical consequences flow from making the conceptual switch to this syllable-centered view? One immediate consequence is a rejection of the assumption that every syllabic position is necessarily occupied by a segment; there may be syllabic positions without any associated segmental content. (Harris and Gussman 1998)

One direction in which this view has been developed has been the pursuit of the hypothesis that there are no codas at all in phonology, and that all syllables are of the form CV.

From this perspective, much of the work of the phonological analysis turns to accounting for where empty vowel positions may occur, for the theory demands that there should be many of them. There is a natural similarity between this kind of phonological analysis and government-binding syntax, in the sense that both require positing a surprising number of unfilled terminal nodes, and in both, a large part of the formal account of what is grammatical and what is not is largely an account of where null nodes may appear. This perspective has been developed by a number of European phonologists, notably Kaye, Lowenstamm, Vergnaud, and Charette, to mention just four, and see Chapter 16 in this book. Considerable

care is necessary in developing a theory of syllables in this direction, if only because it is tempting to think that something has been explained when it has been labeled by a hidden variable: if, in a given language, word-final consonant clusters are permitted that match the cluster possibilities in syllable onset, one must be circumspect in determining just how much explanation is achieved by positing an abstract, silent vowel at the end of the word: the theoretical savings are no greater than what it would take to express the same constraint *without* positing the abstract syllable.[11]

1.13 Derived Sonority

One line of work has developed the idea that one must distinguish between the inherent sonority of a segment and its sonority in a given context (Goldsmith 1993a; Goldsmith and Larson 1993; Larson 1992; Laks 1995; Tchobanov 2002). This framework of dynamic computational models employs numerical values for sonority, and is embedded within a model that includes a learning algorithm, so that adequate values for sonority can be automatically learned from a phonological sample from any given language.

There are three central ideas in this approach: first, that prosodic prominence takes on values on a numerical scale not restricted to integers; second, that there is a difference between inherent (we might say, underlying) prominence and derived, or contextual, prominence; and third, that a language can identify those elements whose prominence is a peak, that is, a local maximum in a numerical sense. "Prominence" here refers to sonority, when considering models of syllabification, and accent, when looking at models of accentuation.

We will limit our discussion to the former case. Thus we must compute for each segment, or timing tier unit, its sonority level, which will be a combination of its inherent sonority and effects that impinge on the unit from the context it finds itself in. The effects divide into two sorts: those that are specific to end units (corresponding to appendix effects: see Section 2.1) in a word, and those that result from the influence of neighboring segments. In the following equation, we define the sonority of the i^{th} segment at time t+1 as the sum of four terms: the segment's inherent sonority, the activation that may be due to a possible edge effect, and a weighted sum of the sonorities of the segments immediately to the left and to the right. The variables α and β specify the degree to which a segment influences its left- and right-hand neighbors; these values are typically negative, giving rise to a competition between neighboring elements. If we note the activation of the i^{th} unit after t computations as A_t^i, then the operative formula is given in (2).

(2) $$A_{t+1}^i = Inherent(i) + Position(i) + \alpha \times A_t^{i+1} + \beta \times A_t^{i-1}$$

The system reaches effective equilbrium after several iterations, and all segments that are peaks of sonority are (that is, are predicted to be) the nucleus of a syllable. Languages typically put minimum sonority conditions on what elements may be

a syllable nucleus, and violations of those minima are the only way in which illicit syllabification arises within this model.

1.14 Optimality Theory

Optimality Theory early on offered an account of the rough typology of syllable types by proposing two syllable constraints that are specific to syllables, Onset (violated by any syllable not containing a filled onset) and NoCoda (violated by any segment in a coda) (Féry and van de Vijver 2003); these can be usefully compared with Pulgram's Rules of *maximum open syllabicity*, and of *minimal coda*, as discussed above. One could imagine a different set of constraints, couched within an optimality theoretic framework, that would account for syllabification, but most work to date has assumed some version of these constraints, or constraint families.

If faithfulness constraints (Dep, Max) are ranked lower than the syllable constraints, then a language will use either epenthesis or deletion to ensure that surface forms are of the form CVCVCV. If Onset is ranked higher than the faithfulness constraints, which are in turn ranked higher than NoCoda, then some strategy, such as consonant epenthesis, will emerge to provide a consonant to precede any vowel that is not preceded by a consonant. If the faithfulness constraints outrank the syllable constraints, then codas, coda clusters, and onset clusters may emerge, if the lexicon and the morphology provide such circumstances. Such an account uses cover terms such as NoCoda, which can be viewed either as promissory notes, or better, as implicit hooks into whatever theory of phonological representation one chooses to use, provided it permits access to coda-labeling as such.[12]

1.15 Must We Choose Between Sonority and Constituency?

I have tried to emphasize in the discussion to this point that there are at least two different pictures of what the syllable is that have evolved, one focusing on the notion of sonority, and the other concerned with constituency structure at and below the level of the syllable. In the next section, we will turn to a third view, the one focusing on patterns of first, or possibly second, order restrictions on segment sequences.

Perhaps all scholars would agree that the core phenomenon lying behind the concept of the syllable is the fact that we can divide the inventory of a language's sounds into two sets, the vowels and the consonants, each of which share a number of articulatory and acoustic properties, *and* that there is a strong tendency for utterances to produce sounds successively, first one from one of these groups and then one from the other group. But things are not really so simple; there are recurring restrictions on what sequences of consonants may occur, but fewer restrictions on what vowel-consonant, or consonant-vowel, sequences may occur. The notion of sonority emerges as soon as we note that we often find a sequence X-Y that occurs at the beginning of a syllable being matched against not the

sequence X-Y, but rather the sequence Y-X, at the end of the syllable. It is this purely structural discovery of a mirror-image of segment sequence possibilities, plus a desire to associate the directionality of a permitted linear sequence of segments with a physical dimension, that brings us to the notion (or *a* notion) of sonority, because what we describe as increasing during the first part of a syllable, and decreasing during the second, *is* sonority.

Must we *choose* between such a sonority-based view of the syllable and a constituency-based view? One possible answer to the question that begins this section is "no," on the grounds that many analyses have been offered that establish explicit connections between the sonority of the segments present and the syllable constituents (onset, coda, nucleus) to which these segments are assigned, and those analyses are not self-contradictory in any obvious way; various OT perspectives, as we have seen, are examples of such analyses. In a recent survey, Zec (2003) sums up her view of the optimality theoretic perspective with the words, "the prosodic constituency is viewed here as a hierarchy of sonority peaks" (p. 125). But there *is* no natural relationship between establishing a height over a sequence of points as in Figure 6.5, on the one hand, and a hierarchical representation, on the other – no natural connection in either direction. To make such a connection, one must add principles of one sort or another; one could offer a positive markedness constraint, as Morelli (2003: 359) does:

> Sonority Sequencing Principle: In a syllable, sonority increases toward the peak and decreases toward the margins.

But a thoughtful review of whether we need both conceptions, and how well they can co-exist, leaves us with some unanswered questions, notably about the notion of the syllable as an example of constituent structure.

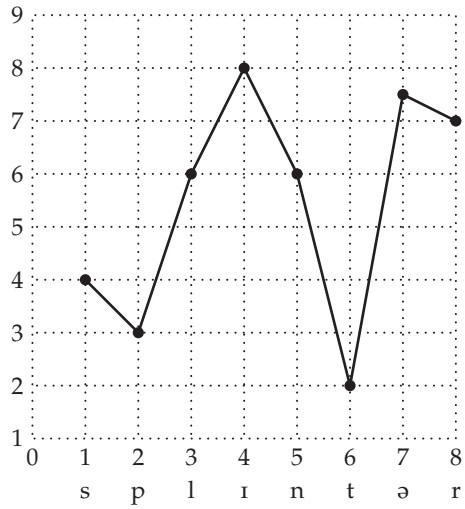


Figure 6.5 A sonority curve.

In the context of syntactic analysis, the notion of immediate constituent was outlined in Bloomfield (1933), and developed by Wells (1947). As Percival (1976) has discussed in detail, analysis in immediate constituents was understood as an alternative to a word-oriented notion of syntax, a view according to which a verb took a subject *noun* and object *noun*, rather than *noun phrase*, though the noun that was the subject or object could have other words modifying it in turn. The development of the theory of immediate constituents was based on two central observations. First, the structure was *hierarchical*, in that if we take a grammatical sentence like *the turtle saw the horse*, we can expand either the subject or the object into indefinitely large constituents (by adding relative clauses and the like). Second, when there are grammatical dependencies between nearby words or constituents, more often than not the dependencies can be analyzed as holding between adjacent constituents that are themselves the immediate constituents of a larger constituent. In French, the choice of determine *le* or *la* "the" depends on the gender of the following noun, and its distribution is best described as being the first of two immediate constituents that form the noun phrase in French.

If we turn then to phonology, we find nothing corresponding perfectly to the way in which sentences are found embedded within one another in syntax; if we find hierarchy, it is by way of a sequence of essentially different kinds of constituents (syllable, foot, word, and various larger phrases), rather than the recursive structure we find in syntax. We find some distributional dependencies, but the dependencies are at least as different from those found in syntax as they are similar to them; we return to this question shortly.

What is most important about constituent structure in syntax, then, does not arise in phonology; and the best-grounded fact about the syllable – that it describes an interval of rising, then falling sonority – has no natural counterpart in syntax; there is nothing which rises and falls in a syntactic sentence, even if it is of the shape SVO (i.e. *Subject-Verb-Object*) let alone if it is VSO or SOV or anything else, as there is in a phonological syllable.

1.16 Phonotactics: Patterns in Sequences

One of the goals of the phonological analysis of a language is to determine, once we have established the set Σ of underlying segments of the language, what sequences of such segments are permitted. What sequences – that is, what subsets of Σ^* – are found in the language, and what subsets of Σ^* are *not* found? The phonologist's goal is to offer meaningful generalizations about the answers to both of those questions.

Phonologists have long hoped to simplify the overall phonological description of possible words by viewing words as being built up out of smaller phonological units, notably the syllable. Pike (1947) emphasized the importance of the syllable in dealing with this observation, noting that the syllable was:

> the basic structural unit which serves best as a point of reference for describing the distribution of the phonemes in the language in question. (p. 144)

Not all phonotactics involve reference to the notion of syllable, however. Scholars have argued that at least some phonotactic statements are best viewed outside the purview of syllables (Lamontagne 1993; Blevins 2003; Kabak and Idsardi 2007). Blevins notes cases, for example, where a neutralization of laryngeal features on a consonant occurs before obstruents regardless of whether the two segments are in the same syllable or not. She suggests, in a similar vein, that in systems which require homorganicity of coda nasals to following obstruents, the restriction should be analyzed as a negative filter against a nasal that is specified for point of articulation and followed by any consonant (i.e. a negative filter against *[*nasal, PLACE*] *Obstruent*) (p. 379), though we should recall Bybee's argument *for* the syllable, mentioned above, that the rule of nasal assimilation occurs in a fashion that is better treated by saying that the environment is across syllable-boundary.[13]

Accepting (as we have throughout this chapter) the simplification that consists of analyzing phonological representations as sequences of segments, we define the possible syllables of a language as a finite set σ of sub-strings of Σ^*, and we define a phonological word as a sequence of elements of σ. This touch of formalism is intended to make natural the following question: what is the right way to characterize the set σ?

It is possible to focus one's theoretical account on a compact set of statements regarding what segments may follow what segments: in effect, to offer a finite state automaton as our model, where each edge of the automaton generates a segment of the language. On this account, we have a formal device that generates, as we might say, "left-to-right," – that is, a directed graph, whose paths correspond to all and only the possible sequences within a syllable. Possible clusters of three or more consonants have been analyzed in several languages as consisting of exactly those that can be analyzed as overlapping pairs of permitted two-consonant sequences (that is, $C_1C_2C_3$ can exist if and only if C_1C_2 and C_2C_3 can independently exist), and it is natural to interpret this kind of analysis as employing a finite state model.[14]

In studies of finite state automata (FSA), it is common to consider two ways of thinking of a string as being generated by a particular path through such an automaton from beginning (#) to end (#): either *states* are associated with symbols which may be emitted when a path goes through a state, or else *edges* (from one state to another) are associated with symbols which may be emitted when a path follows a particular edge. In Figures 6.6–6.11, I have illustrated the former style of FSA for languages with simple syllable patterns, while in Figures 6.12 and 6.13, we see the latter style of FSA, to which we now turn.

A very early effort at describing in detail the sequential structure of English syllables was made by Benjamin Lee Whorf (1940), whose goal was to illustrate the complexity of the implicit knowledge of any native speaker of English. He presented it in a format which is similar to, but by no means identical to, a finite state automaton (and he was working before the notion had entered the literature). I have therefore modified it a bit to look more like a familiar finite-state diagram. The part covering the syllable onset is given in Figure 6.12; all paths lead to a

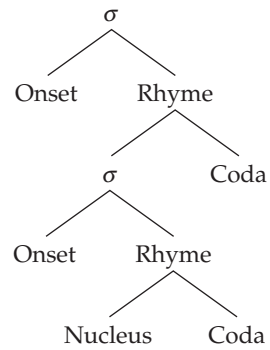


Figure 6.6 A monstrosity? Recursive constituent structure.

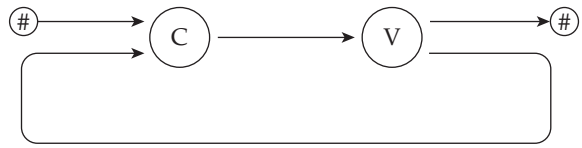


Figure 6.7 CVC syllables.

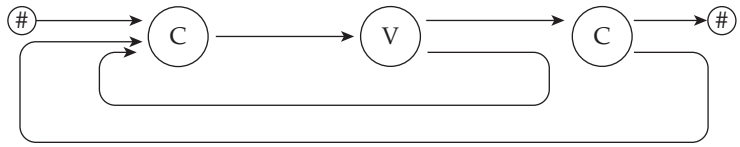


Figure 6.8 CV(C) syllables.

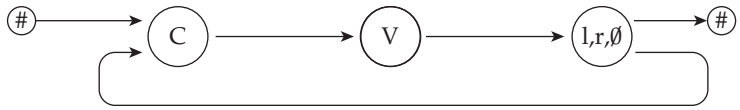


Figure 6.9 CV(L) syllables.

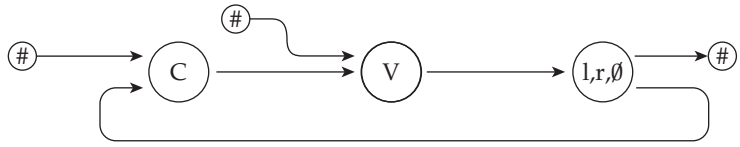


Figure 6.10 (C)V(L) syllables.

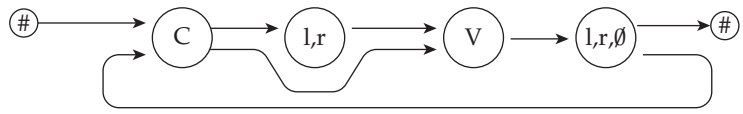


Figure 6.11 C(L)V(L) syllables.

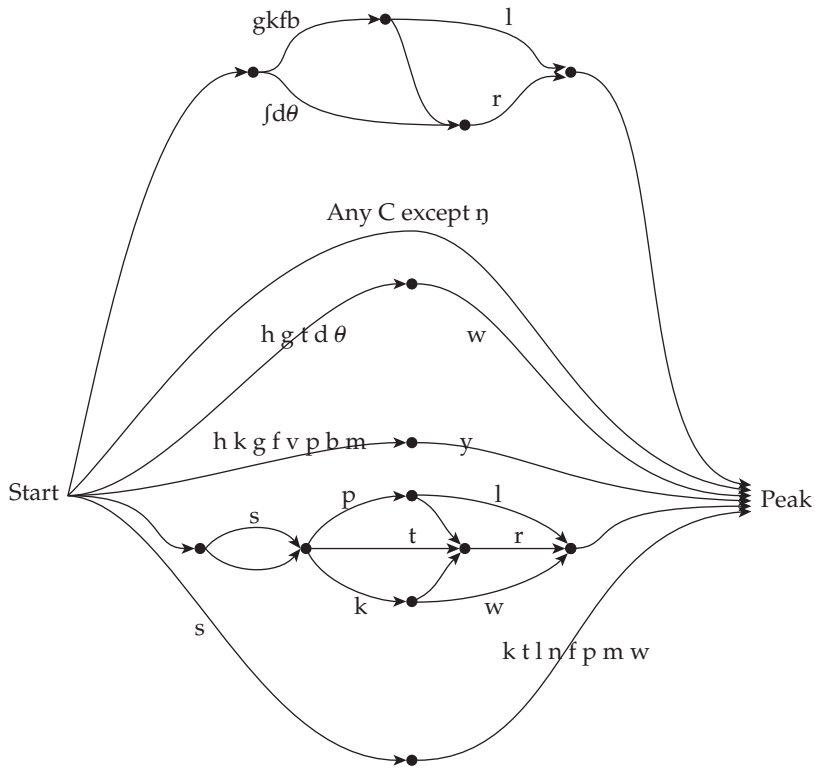


Figure 6.12 Whorf's onset.

single state (though there is an added complexity involving the *yu* diphthong), and this convergence explicitly represents the lack of dependencies between the choice of onset and the following nuclear vowel. In Figure 6.13, his rhyme is represented in a similar fashion. The intention is for there to be a one-to-one association between paths in the graph and possible onsets in English (though the graph fails to generate the cluster *kl*, apparently an overlooked flaw). Each edge is associated with a set of segments, and one such symbol is generated when taking an edge from one state to the next, if there are segments associated with the edge; if there are none, then passage along that edge does not contribute any symbol. Several edges are associated with the null set of symbols (equivalently,

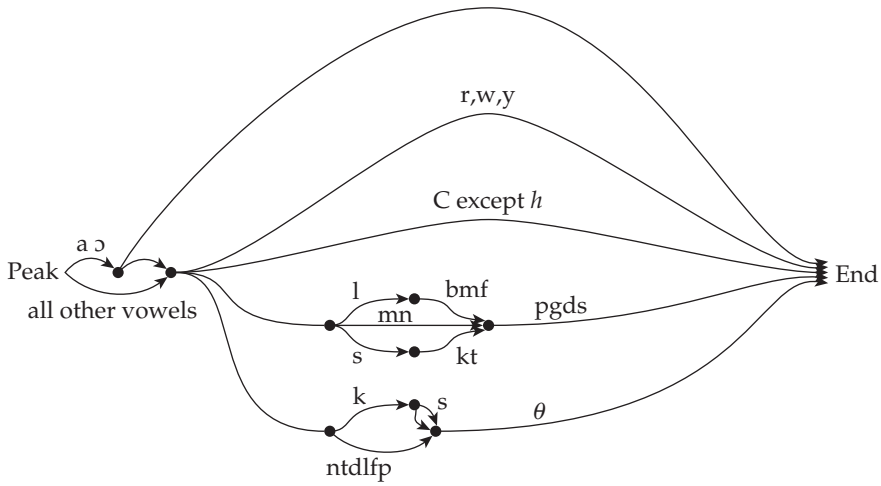


Figure 6.13 Whorf's rhyme.

are associated with the null symbol), though this is essentially a characteristic that I have had to add in redoing Whorf's notation. In one place, an unusual notation is necessary: one edge is labeled as generating any consonant *other* than the velar nasal *ŋ*.

An analysis of this sort does not indicate any hierarchical structure, but it obviously is one which contains a lot of structure: the structure inheres in the statement of permissible sequences, rather than in the representation of any particular sequence. In modern parlance, it is *non-deterministic*, in the sense that from several states, more than one path leading from the node can generate the same symbol (for example, a syllable-initial *g* can be generated by taking the top-most path, or by taking the third highest path, the one that generates any single consonant other than ŋ). Such formal devices are in some ways better suited to express permissible sequences; the range of their abilities is different from that of phrase-structure rules. In addition, a model of this sort can easily be made probabilistic, and to illustrate a first-order Markov model. This allows us to easily indicate the difference between the probability of (for example) a syllable-initial *p* and a *p* that immediately follows an *s*.

1.17 Onsets, Codas, and Word-appendices

Phonologists have long hoped to simplify the overall phonological description of possible *words* by viewing words as being built up out of smaller phonological units, either syllables or, in more recent work, feet (which are, in turn, composed of syllables). After all, a rough account of possible sequences of sounds in words can often be formulated by saying that a phonotactically admissible word is one

that can be analyzed as a sequence of phonotactically admissible syllables. On this view, there should be perfect agreement between what consonant sequences can occur word-initially and syllable-initially, just as there should be perfect agreement between what consonant sequences can occur word-finally and syllable-finally.

However, it has been known for a long time that this approach is much too simple to account for the facts of language. The effort to reduce well-formed words to sequences of well-formed syllables runs into serious trouble when either of the following holds:

(i) the set of consonant strings (*C-strings*) that appear between vowels turns out not to be identical to the set of all sequences that we obtain by concatenating a word-final C-string plus a word-initial C-string; or
(ii) if we have some clear way of determining where the break is between syllables – let us suppose that we have an inter-vocalic sequence C and we can determine that it is broken between syllables as C_1 and C_2, where $C = C_1C_2$ – and either C_1 does not occur word-finally, or C_2 does not appear word-initially, or both.

Both of these situations have long been known to exist. Languages exist, for example, in which no more than two consonants appear word-initially, and no word-final consonants, but word-internally, sequences of three consonants are found (we will see examples below). Inconsistencies may arise in the other direction: words may begin or end with sequences that are not permitted (as onsets or codas, respectively) word-internally. Similarly, while Dutch permits *str*, *spl*, and *spr* clusters word-initially, such clusters are split *s-tr*, *s-pl*, and *s-tr* when they occur word-internally, to judge from the laxing that occurs in closed syllables in the case of words such as *mistral*, *esplanade*, and *Castro* (Trommelen 1983; van Oostendorp 1995). Or again, word-initial clusters such as are found in Dutch *gnoom*, *slaaf* and *tjiftjaf* are syllabified in separate syllables when the sequences are found word-internally (van der Hulst and Ritter 1999b).

These problems have been attacked in several ways. If word-initial combinations appear to be more numerous than other syllable-onset combinations, then an extra word-initial position could be proposed, and similarly, if word-final sequences are richer than syllable coda sequences, extra word-final consonant positions could be proposed. These positions are often called *appendices*. (In addition, other considerations may be brought to bear, notably the difference between what consonantal material may appear in the onset and coda of a stressed syllable and of an unstressed syllable.)

The strongest argument that has been made for approaches that employ language-specific appendices is based on the observation that word-final appendices are also inert or invisible with respect to measures of syllable weight, which in turn are relevant to stress assignment (see Chapter 5). It is also interesting to note that word-initial and word-final appendices often violate sonority sequencing generalizations that hold word-internally, though this observation may carry somewhat less weight (van der Hulst and Ritter 1999b: 16). In addition, it is found that in some

(perhaps most) languages with closed syllable shortening, with only a single consonant permitted in the coda of a non-final syllable and two consonants permitted in word-final position, a long vowel is permitted in a word-final syllable just in case it is followed by one, and not by two, consonants – all of which suggests that the final consonant in such a system is parsed as a word-appendix rather than the coda consonant of the final syllable. See Vaux and Wolfe (2009) for a recent in-depth look at this subject.

In some cases, there appears to be a clear link between appendix-like behavior and morphological status. Of this, English provides a simple example: except in morphologically specific cases, everywhere in English a coda nasal will be homorganic with a following obstruent in the same coda, as in *bank* and *slump*. But before the regular verbal suffix *-d* and the plural suffix *-z*, we find violations of this (*flamed, banged, flames, bangs*).

This general situation was noted as early as 1952, in an important survey (Fischer-Jørgensen 1952), where Fischer-Jørgensen noted that there are languages in which "some medial clusters cannot be dissolved into actually occurring final and initial clusters" (she cites Italian, Totonaco, Chontal, Yuma, and Kutenai), adding that

> it is evident that the phenomenon is not rare. But generally these cases are exceptions, even within the system of the language in question . . . But there are very extreme cases of this phenomenon . . . Finnish constitutes a good example. In Finnish the only consonants admitted finally are *n*, *r*, *l*, *t*, *s* and initially genuine Finnish words have only one consonant; but medially a great diversity of clusters is found, e.g. *ks*, *rst*, *mp*, etc. (p. 306)

She notes the even starker example of Keresan, in which words obligatorily begin and end with consonants, but CVCVC and CVCVCVC are fine word-patterns, leading the analyst to the conclusion that while CV syllables are permitted word-internally, such an open syllable cannot appear word-finally.

Some work has tried to tackle the analysis of word appendices as supernumerary onsets (as word-final appendices) or codas (as word-initial appendices), a view that is very close to the proposal that these edge effects are the result of syllables that are degenerate, in the sense that they do not contain a vowel, something that is typically taken to be an *sine qua non* for a syllable, after all. Perhaps the earliest example of this is cited by van der Hulst: he notes that Kuryłowicz (1952) treats this situation as involving a stranded onset. On word-peripheral clusters, see, for example, on Dutch: Trommelen (1983), van der Hulst (1984); on Polish: Cyran and Gussman (1999), Rubach (1990), Rubach and Booij (1990a), Davis (1990); a good survey appears in Törkenczy and Siptar (1999). Kiparsky (2003b) offers an interesting account of striking differences in modern Arabic dialects based on differences in the ranking of a constraint requiring moras to be licensed by syllables; where the constraint is violated, structures are found in which phonological material appears despite it going well beyond the range of possibilities permitted by Arabic core syllables.

It has long been noted that fewer consonants and consonant sequences are permitted in codas than in onsets, though this observation has eluded precise formulation. One of the challenges to dealing adequately with this phenomenon arises from the fact that in many cases, it is not so much a segment *type* that is excluded from the coda as a segment *contrast*. In what seems to be positive terms, we can say that the coda is a position of neutralization; in negative terms, we can say that a certain segment type cannot appear in the coda unless it is the result of a generalization that expresses a neutralization! The central example of such phenomena is the appearance in many languages of geminate consonants intervocalically: in a form like Italian *detto* (Figure 6.2), the coda of the first syllable has a [t] in it, but this is possible only because the following consonant, in the syllable onset, is a [t] – which is to say, a geminate can give rise to a structure in which an obstruent appears in a syllable coda, even in a language which does not permit, in non-geminate cases, a coda obstruent. This kind of situation has been described in terms of the logic of licensing; see Itô (1986), and, for a slightly different perspective, Goldsmith (1989).

2 Syllable-based Alternations

It is a commonplace to find in the phonological literature descriptions of phonological alternations of consonants in which the crucial context is the syllable position of the consonant: it is realized in one way when in the syllable onset, and another way when in the coda. Things are often not *quite* that simple, but that serves as a central focus of a wide range of phenomena, as illustrated below.

2.1 Spanish s

The behavior of /s/ in New World Spanish dialects illustrates a common pattern in which a consonant is realized differently in an onset and in a coda. /s/ in onset position is realized as [s], but in many dialects, /s/ in coda position is realized as [h], which is to say, as an [s] whose oral frication is removed. This phenomenon, often referred to as *aspiration*, is widespread in New World Spanish, and attested in Peninsular Spanish. In dialects with aspiration, *esta* "this" is [ɛhta]; there is considerable dialectal varation with regard to the behavior of word-final /s/.

2.2 French loi de position

Most dialects of French have six oral mid vowels: ɛ, e, o, ɔ, ø, œ. Of these, three are open (ɛ, ɔ, œ) and three are close (e, o, ø). There are conditions on their distribution, however. A close mid vowel cannot appear in a closed syllable. For the pair (e, ɛ), we see this effect in morphophonemic alternations as well as in many loanwords from English, where a tense and diphthongized vowel in English [e^{y}] is borrowed as [ɛ] in closed syllables, e.g. [tɛk] English 'teak', [mɛl] 'e-mail', [stɛk] English 'steak', [kɛk] English 'cake', and in truncated forms in contemporary

speech, as in [tidɛʒ] from [pətideʒøne] 'breakfast', or [agrɛg] from [agregasjõ] 'teaching certification'. In Québécois French, this relationship is extended to high vowels as well: (i, y, u) appear in open syllables, and (ɪ, ʏ, ʊ) appear in closed syllables, as well as in syllables to the left of a closed syllable to which regressive vowel harmony has applied (e.g. [dzɪfsɪl] 'difficult'). This is attested in pairs of related words, such as [ptsi] [ptsɪt] 'small' masc., fem.

2.3 *Flapping in English*

Trager and Bloch's classic, but controversial, analysis of English phonology (1941) argues that in the case of word-internal stressed-unstressed sequences, such as *bidding, bedding, padding, nodding, budding*, and *pudding*, each of "the six short vowels is followed by an ambisyllabic voiced stop" (p. 233), and in a curious appendix to a very brief note in *Language* (see Eliason 1942), they support their view of the proper syllabification of English words such as *hitting*: "the division here is not before the medial consonant and not after it – in short, that the consonant is ambisyllabic, and that the division occurs, if at all, within the consonant itself" (p. 146). Kahn (1976) developed an analysis of flapping in American English which strongly supports their analysis.[15] The following analysis departs from Kahn's analysis in some of the specifics, but follows it in overall construction.[16]

One of the most striking characteristics of American English is the widespread appearance of the coronal flap [ɾ] as a realization of /t/ as well as /d/, and the principles that lie behind the distributional generalizations of the flap have led linguists to view the conditioning to be based on syllable affiliation: a coronal stop /t,d/ is realized as a flap if it is simultaneously a member of the coda of one syllable and the onset of the next syllable.

It is not hard to find phonological descriptions of American flapping which state a generalization along these lines: a coronal stop is realized as a flap when it is immediately preceded by a stressed vowel, and immediately followed by an unstressed vowel, in *Italy*. While this is true, it is only a small part of the story. In the real description of American flapping, it is first of all necessary to distinguish the conditions under which word-internal /t/s are flaps from the conditions under which word-initial /t/s and word-final/t/s are. Consider first the case of strictly intervocalic, word-internal /t/s, where the facts roughly follow the description just given. The context ṷ–v̆ mentioned there is, in fact, a position in which flap obligatorily appears: for example, *Italy* [íɾəlĭ]. There are three other strictly intervocalic contexts to consider: ṷ–ṷ, v̆–ṷ, and v̆–v̆. In the first two, we do not get a flap at all; it is not possible in words such as *bótòx, détàil, rétàil, látèx, Útàh*; or *Ĭtálian, ăttáin*, etc. In the third case, where the /t/ is surrounded by unstressed vowels, as in *sanity* or *opacity*, both flapped and unflapped variants are possible (they are equally acceptable to this writer). The generalization does not change (here as elsewhere) when we extend the context to include a preceding *r*; *parting* and *potting* have flaps just the same. The same is not true of other sonorants: a /t/ will not flap after /l/; we have *faulty* with no flap possible, for example. (There is a complication when a syllabic *n* follows the /t/, as in *Latin*, which we will ignore here.)

Word-final /t/ may always be pronounced [t˺], a glottalized and unreleased [t], associated with at least a weak phrasal boundary immediately following, but in connected speech, when the following word begins with a vowel, regardless of whether the vowel is stressed *or* unstressed, a flap is found – and this is true whether the vowel preceding the /t/ is stressed or unstressed. Examples of the four cases, where the [ɾ] realizes a /t/: *Gé[ɾ] ŭs out of here! Gè[ɾ] óut of here! A lockĕ[ɾ] ŏf hair. The rabbĭ[ɾ] áte the carrot*. In these cases, an empty onset attracts a preceding consonant, even if the consonant is "already" syllabified. The result of this is that the /t/ is ambisyllabic, and ambisyllabic /t/ś are flapped. This is not a case of maximizing onsets; a /t/ does not resyllabify before an *rV* sequence, for example *(the ha[t˺] ripped*, with no possibility of the /t/ being part of an onset-cluster).

The third case, that of the word-initial /t/, depends, curiously enough, on the particular word in which it is found. If the word is *to, today, tonight, tomorrow*, or *together*, then we find one behavior, which I will temporarily refer to as *to*-behavior; if the word is any other (for example, *tomato, tuba, Topeka, topology, Thomas, taste*), we find a different behavior. Furthermore, the realization is largely independent of whether the preceding vowel is stressed or not, and largely independent of whether the following syllable is stressed or not. There is no flap in *the tomato, a tenacious opponent*, or *a topology*, where the /t/ is in an unstressed syllable, nor in *the total* or *the toast*, where /t/ is in a stressed syllable. Thus this case is entirely different from either the word-internal or the word-final case.

However, in the case of the handful of words based historically on the preposition *to* (*to*, *today*, *tonight*, *tomorrow*, and *together*), the facts are different. In each case, flapping is possible (indeed, preferred) when the preceding word is vowel-final: *Go [ɾ]o sleep! How'd it go [ɾ]oday?*, etc.

A natural way to interpret this data involves two passes of syllabification. The first applies word-internally, syllabifying a segment to an immediately following vowel, regardless of stress, and a rule that adds a syllable link between any open syllable and an immediately following consonant in the word (optionally if the syllable is unstressed, and obligatorily if the syllable is stressed). This results in an ambisyllabic consonant. At the phrase level, only one rule is operative: a word-final consonant adds an affiliation to a following syllable σ if σ begins with a vowel in the same phonological phrase. That rule also results in an ambisyllabic consonant. Given these two rules, we may say that any, and only, ambisyllabic /t,d/ is realized as a flap [ɾ]. The *to-* initial words that we noted above are all cliticized to the word that precedes, in the sense that it is treated as a single phonological word with what precedes it.

3 Conclusions

What can we conclude about the syllable, in the light of the studies that we have reviewed? There are repeating patterns of sequences of sounds in language, and these patterns define the syllables of various languages, and these patterns lie at

the base of many, or all, prosodic phenomena. But how are these repeating patterns best described, and how are they best explained? We have an embarrassment of riches in facing both of these questions. In this final section, we will reflect a bit on how this is so, and what we might do about it.

Phonology, as a field, is still struggling to deal with the consequences of the development of the phoneme, which is at the same time its greatest achievement. By "the phoneme," we mean the abstract characterization of a set of sounds in a language which unifies all of the sounds into a relatively small inventory of elements which are then used to define contrasting morphemes and words. This insight is the beginning of all work in phonology. Yet at the same moment, two other types of analysis – analysis into syllables, and analysis into features, or in short, analysis into units both larger and smaller than the phoneme – are crucial for any descriptive account of the phonology of a language.

On the one hand, languages never offer the unconditional sequence of any phoneme followed by any phoneme: local conditions of dependence are present everywhere. Using the phrase "local dependence" suggests that a first-order Markov model might be a good model of phoneme occurrences: segments do indeed care very much what their neighbors are, so to speak. But mapping out the conditional probabilities of each phoneme, based only on what phoneme precedes, fails to capture the just slightly larger generalization that not just lurks, but looms, behind the data: to wit, that while many languages permit sequences of two consonants, very many exclude sequences of three. We could expand our vision to a second-order Markov model, allowing each phoneme's options to be limited by the two phonemes that precede it, but we would be losing sight of the bigger generalizations. That is, if there are p different phonemes in the language, there are p^3 different parameters that need to be specified for a second-order Markov model: each phoneme's probability after each pair of phonemes would need to be specified. But any study of a real phonology shows us that only a small portion of the universe of p^3 possibilities has a chance of being utilized by a natural language phonology, *because* there are generalizations just slightly larger in scope.

These generalizations involve what we call the syllable. But how should these generalizations be modeled and formalized? We have seen three major traditions over the course of this chapter, the syntax-based immediate constituent approach, the sonority approach, and the finite-state approach. The first specifies constituents of structure and utilizes phrase-structure rules to describe possible sequences, the second maps each element of the phonemic inventory to the real numbers, and then reconstructs conditions on possible numerical sequences, such as limiting which phonemes can appear at local peaks of sonority, while the third focuses its theoretical capability on a statement of what sequences are permitted in a given language.

On the other hand, syllabification is not simply an effect, of which the sounds are the cause: quite to the contrary, the choice of phoneme in some cases, and the choice of allophone in a very large number of cases, is determined by the location of a sound in the larger prosodic stream. Of this, the most striking special case is the difference in the realizations of consonants in syllable onset and in syllable coda.

And yet clear evidence of constituent structure in phonology is notoriously difficult to establish, certainly compared to the ease with which we can determine that choice of the allophone (realization of a phoneme) is conditioned by the immediately following phoneme, and compared to the ease with which we can distinguish between the characteristics of a consonant in the coda and in the onset of a syllable.

My conjecture is that the syllable is ultimately best regarded as the lowest level (or one of the lowest levels) of rhythmic reoccurrence of possibities in language. Some might want to see this as the reflection of gestures made by the articulatory apparatus, a view that we have not surveyed in this chapter. For myself, I think that such a view analyzes language at the wrong level of abstraction or granularity: the correct level of abstractness for the description of language is higher than that of jaw gestures. Sonority, and the wave-like recurrence of peaks of sonority, seems to me to be the fundamental pattern of syllabification in language.

Studies that explore the consequences of optimality theory for our understanding of the syllable, and vice versa, seem to me to largely miss the point that we have discussed in this chapter, and in a sense that should not be surprising, in view of what optimality theory is: it is a theory of constraint interaction, rather than a theory of phonological representation, and it is not fundamentally a theory of how the constraints (appropriate for natural language phonology) should be formulated, even if some phonologists have implicitly, or on occasion explicitly, made some suggestions along such lines. Optimality Theory is perfectly consistent with any of the three views described here.

We began this chapter with a quotation from Ernst Pulgram, and we will end it with another. Pulgram wrote,

> [The syllable] has no function, no *raison d'être*, apart from that of the syllabic segmentation of an utterance. It serves nothing but itself, as it were; it does not serve, immediately like a sign or mediately like a figura, the communicative purpose of a language . . . A syllable is . . . a phonological unit that is, as all linguistic units must be, describable and definable only on its own level of analysis exclusively.[17] (p. 21ff.)

Perhaps that is the best we can do for now. But I think that it is not the last word to be said on the subject. The most important question to answer is how to develop a model that is suited precisely to capture the rhythmic character of syllables, and the striking asymmetries of onset and coda. We have amassed a great deal of knowledge in recent decades that will help us reach that goal eventually.

NOTES

1 I thank Peter Auer, Juliette Blevins, Diane Brentari, Stefano Canalis, Stuart Davis, François Dell, Tracy Alan Hall, Harry van der Hulst, Jonathan Kaye, Paul Newman, Marc van Oostendorp, Jason Riggle, Donca Steriade, Bert Vaux, Ilya Yakubovich, and Alan Yu for comments on an earlier draft.

2 Hyman (1983). Auer (1994) criticizes the view that all languages are syllable-oriented, and explores the difference between syllable languages and word languages.

3 I have just distinguished between phonotactic regularities and conditioning regularities, but both types of regularities describe conditions on what segments may follow one another in a language. The distinction between the two rests on the assumption that it is possible to specify, for a given language, an inventory of underlying sounds categories, the traditional phonemes or underlying segments. In the context of this book, that assumption is not controversial. Phonotactic regularities are then statements as to what sequences of phonemes are permitted in a language, while the other *conditioning* regularities are the statements in a phonology about what phonological elements may cooccur with what; the largest part of this is composed of the rules of allophony, that is, non-neutralizing rules. Syllable position plays a major role in conditioning the realization of an phonological segment. For example,

- The realization of a consonant is frequently different depending on whether it is in a syllable onset or syllable coda.
- The realization of a vowel is frequently different depending on whether there is a consonant in the immediately following coda or not.

4 The title of Sevald *et al.* (1995) gives a good flavor of this perspective: "Syllable structure in speech production: are syllables chunks or schemas?" as does the title of Cutler *et al.* (1986): "The syllable's differing role in the segmentation of French and English."

5 I am indebted to Bernard Laks as well as to Ali Tifrit for bringing the importance of this material to my attention. See Laks (2003) and Tifrit (2005).

6 It is worth bearing in mind that there *was* in fact a frontier of research in syntax at the time, and that notions that many of us take for granted today were being developed and argued about during the 1940s, and also that it was not a *discovery* that the structure of the syllable matches that of the sentence within a constituency-based theory of the syllable, since this theory was specifically created in order to have this appearance. Canalis (2007) discusses the influence of the work by Hjelmslev, with Uldall, in the framework of glossematics, noting that Hjelmslev (1939) should be cited in the development of the constituency view of the syllable. The history of this period needs to integrate Hjelmslev's influential work.

7 See also Haugen (1956), and the position of Clayton (1976), Shibatani (1973), and Hooper (1976).

8 One of the earliest explicit discussions of the significance of identifying *one* of constituents of an immediate constituent as the nucleus or head, and the other as non-head or satellite, is Pittman (1948), though he does not discuss phonology *per se*. A number of phonologists have explored this asymmetry over the last 20 years; some of them have worked in a framework influenced by government phonology, though not all have. This question is discussed in Chapter 16.

9 In fact, Microsoft recently patented this idea, or perhaps just something very close to it; see US Patent 20050203739, granted in September 2005. In a context such as this, the mutual information between two adjacent segments s_1 and s_2 is $log \frac{prob(s_1\ s_2)}{prob(s_1)prob(s_2)}$. See Goldsmith (2007), Goldsmith and Riggle (2012).

10 Pulgram presents these rules as ordered, but this reader gets the impression that his use of ordering is essentially for the purpose of indicating that a later rule has empirical precedence over an earlier one.

11 See the introduction in Hartmann *et al*. (2008) for a general discussion of empty categories.
12 Some offer this account as a success of OT, as in Féry and van de Vijver (2003: 8), where the authors take a different view: "The ability of OT to explain typological patterns as a result of the interaction of markedness and faithfulness constraints is the core of the theory, and it is to a great extent responsible for its success" (p. 8). Others might ask for an account of why the constraints are NoCoda and Onset rather than Coda and NoOnset, or Coda and Onset, or NoCoda and NoOnset.
13 See also Steriade (1999a).
14 See Fischer-Jørgensen (1952), who cites Bjerrum and Hjelmslev, though she gives counter-examples from Russian and Kutenai; see also Clements and Keyser (1983), as well as Hjelmslev (1939) and Butt (1992).
15 Some scholars were unpersuaded that the facts should be described with terms including *ambisyllabicity*; see, for example, Picard (1984), who does not appear to be familiar with the historical depth of this view, viewing it rather as an artifact of Kahn's design.
16 Kahn's analysis crucially involves ambisyllabicity. Such an approach has been challenged by, among others, Kiparksy (1979) and Selkirk (1982); see also Harris and Kaye (1990), Hammond (1997); also Rubach (1996), Jensen (2000) and Picard (1984). Alternatives to the ambisyllabicity approach need to appeal to using feature specifications on a segment to give it a mark indicating its syllabic position earlier in the derivation.
17 Pulgram's remark actually calls to mind Mark Aronoff's recent argument (1994) that what he calls *morphomes* in language have an existence that is, in many cases, for themselves and only for themselves: they are more concrete than morphemes, and play an important role in the morphologies of many languages.

7 Tone: Is it Different?

LARRY M. HYMAN

1 Introduction

Except for a brief period in the late 1970s and early 1980s, tone has generally fallen outside the central concerns of theoretical phonology. During that period, the concepts and formalisms of Goldsmith's (1976a, b) autosegmental approach to tone provided the model to address other aspects of non-linear phonology including vowel harmony (Clements 1976, 1981), nasal harmony (Hyman 1982), and feature geometry (Clements 1985; Sagey 1986). In addition, autosegmental approaches to templatic morphology (McCarthy 1981), reduplication (Marantz 1982), and other aspects of prosodic morphology owed their inspiration to tone, which through the work of Pulleyblank (1986) provided important insights into the developing framework of lexical phonology and morphology (Kiparsky 1982b, 1985; Mohanan 1986). Most generative work prior to and during this period had centered around African tone systems (Leben 1973a; Hyman and Schuh 1974; Goldsmith 1976a; Clements and Ford 1979; Clements and Goldsmith 1984), two notable exceptions being Haraguchi (1979) and Yip (1980), who dealt with the tonal dialectology of Japanese and Chinese, respectively. Finally, Pierrehumbert (1980) developed an influential autosegmental approach to intonation based on English, which was subsequently applied to Japanese (Beckman and Pierrehumbert 1986) and many other languages since. While the autosegmental legacy is still quite alive, tone has not contributed as centrally to subsequent theoretical innovations in phonology. In the case of optimality theory (Prince and Smolensky 1993; McCarthy 2002), there have been some interesting applications, for example Myers' (1997) treatment of the Obligatory Contour Principle (OCP), but theoretical developments have largely been based either on segmental phonology or on stress, syllabification, reduplication, and other aspects of prosodic phonology and morphology.

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

My goal in writing this chapter is twofold. First, I propose to cover some of the aforementioned contributions that tone has made to phonological theory. Second, I wish to show that there is still much more for phonologists and others to learn from tone. I suggest that linguists should be very concerned about tone, for at least three reasons:

(i) Tone systems are found in approximately 50% of the languages of the world. The greatest concentrations of tone languages are found in Sub-Saharan Africa, East and Southeast Asia, South-central Mexico, and parts of Amazonia and New Guinea. While we have had access to information about the first three areas for some time, comparatively little has been available on tone in Amazonia and New Guinea until recently. The emerging picture is that these tone systems have interesting and diverse properties which complement the already varied African, Asian, and Mexican systems. The result is an extraordinary richness and a potential gold mine for future investigations.

(ii) The study of tone has influenced the history of phonology and promises to contribute further to our understanding of language in general, particularly interface issues. For instance, some of the most detailed and influential studies concerning the syntax-phonology interface have drawn from tonal alternations applying at the phrase level (Clements 1978; Chen 1987). More recently, several meetings have brought together scholars interested in the relation between tone, phrasal accent, and intonation, especially in the languages of Europe and East Asia (Germanic, Slavic, Basque, Chinese, Japanese, Korean, etc.), areas covered in some detail by Ladd (1996) and Gussenhoven (2004). Three collections of note are van der Hulst (1999), Jun (2005), and Riad and Gussenhoven (2007). Cutting across the components of grammar, both abstract and instrumental research have been concerned with how focus and other aspects of information structure, often marked by stress or intonation, are realized when there is a competing tone system. Some of this work has shown that focus is not necessarily prosodically marked in certain languages with tone (Downing 2007).

(iii) Tone systems have properties which SURPASS segmental and metrical systems. In Section 6 I conclude that tone can do everything that segmental and metrical phonology can do, but that the reverse is not true. This is especially true of the long-distance effects that tone exhibits both within and across words, as when the tone of one word migrates several syllables or words to its right. Since some tonal phenomena have no segmental or stress analogues, anyone who is interested in the outer limits of what is possible in phonology would thus be well-served to understand how tone systems work.

Despite the widespread occurrence of tone in the world's languages and the important contributions it has already made to our understanding of phonology and its interfaces, the lack of familiarity of some scholars with tone has allowed certain old misconceptions to persist. The rest of the chapter is organized as follows. In Section 2 I begin by defining tone and characterizing these persistent misconceptions. In Section 3 I discuss the autosegmental insight into tone. Section 4 deals with the question of whether tone has different properties from other phonology, either quantitatively (4.1) or qualitatively (4.2). Section 5

considers the issue of whether tone should sometimes be analyzed in accentual terms. Throughout the discussion the focus will be on the question of what tone has to teach phonologists, and therefore, crucially, whether there are phonological properties that are found only in tone systems. The conclusion in Section 6 summarizes the key points of the preceding sections, concluding that the capabilities of tone do surpass those of either segmental or metrical phonology.

2 Defining Tone: Three Misconceptions

A logical place to begin is by raising an old, but most essential question, "What is tone?" How do we know if a language has tone? Two early definitions state that a language with tone is a language

(1) a. ". . . having significant, contrastive, but relative pitch on each syllable" (Pike 1948: 3)
b. . . . in which both pitch phonemes and segmental phonemes enter into the composition of at least some morphemes. (Welmers 1959: 2)

While Pike originally saw tone as a contrastive feature on each syllable or other tone-bearing unit (TBU), Welmers' definition insists on the MORPHOLOGICAL nature of tone: tone is not a property of syllables, as expressed by Pike, but rather of morphemes. Welmers correctly pointed out that not all morphemes need to have a tone – some may be toneless. Similarly, not all morphemes need to have a TBU – they may be "tonal morphemes" (Section 2, Section 3). We will see that there is great advantage in approaching tone from this morphological perspective. However, it is useful to update and slightly modify Welmers' definition as in (2), and say that a language with tone is a language

(2) . . . in which an indication of pitch enters into the lexical realization of at least some morphemes. (Hyman 2001c: 1368; Hyman 2006: 229)

This statement defines tone in terms of individual morphemes whose tones may combine (and interact) in forming words. Since word-level tones may be assigned by rule, "lexical realization" refers to the output of lexical phonology, not necessarily underlying representations. For example, it could be argued that the Somali noun roots in (3) are underlyingly toneless:

(3)

	root	masculine		feminine	
a.	/inan/	ínan	'boy'	inán	'girl'
	/naʕas/	náʕas	'stupid man'	naʕás	'stupid woman'
	/goray/	góray	'male ostrich'	goráy	'female ostrich'
b.	/darmaan/	darmáan	'colt'	darmaán	'filly'
	/ʕeesaan/	ʕeesáan	'young he-goat'	ʕeesaán	'young she-goat'
	/dameer/	daméer	'he-donkey'	dameér	'she-donkey'

For these nouns, the masculine morpheme assigns a high (H) tone (´) to the penultimate vowel, while the feminine morpheme assigns an H tone to the final vowel (Hyman 1981; Saeed 1999). Toneless vowels which follow an H are realized low (L), while vowels which precede an H tone are realized mid (M). The contrast in (3a) is thus between [H-L] and [M-H], while in the [M-HL] and [M-MH] contrast in (3b), the last syllable contains a long vowel with either falling or rising pitch. Somali also illustrates the importance of the wording "indication of pitch" in the definition in (2), which is taken to mean tone features or any other analytical device whose only function is to characterize pitch. In the above analysis, an H tone is assigned to one of the last two vowels of the noun. Although an analyst may prefer to assign a diacritic accent (*) to the relevant vowel instead, the diacritic device would still be marking only H tone, which in fact may be entirely absent on a word, for example on both the subject noun and verb in the utterance *inan wáa dhaʕay* 'a boy fell'. The definition in (2) would thus include Somali, but would exclude languages where all indications of pitch are introduced post-lexically, that is, at the phrase or utterance level.

While it is possible to view the above Somali roots as underlyingly toneless, and the gender markers as tonal morphemes devoid of a TBU, the most common situation is for morphemes to consist of both segmental and tonal features. This is seen in the minimal pairs, triplets, quadruplets, and quintuplets in (4)–(7) showing examples contrasting two, three, four, and five contrastive tone heights, respectively:

(4) Two levels: Dadibi (Papua New Guinea) (MacDonald and MacDonald 1974: 151)

a.	L (low) tone :	wà	'string bag'	nà	'shoulder'
b.	H (high) tone :	wá	'edible greens'	ná	'aunt'

(5) Three levels: Nupe (Nigeria) (Banfield 1914)

a.	L (low) tone :	bà	'to count'	wà	'to scratch'
b.	M (mid) tone :	bā	'to cut'	wā	'to extract'
c.	H (high) tone :	bá	'to be bitter'	wá	'to want'

(6) Four levels: Chatino (Yaitepec) (Mexico) (McKaughan 1954: 27)

L (low) tone	Lower mid tone	Higher mid tone	High (H) tone
kn̰	kn̻	kn̄	kń
'dove'	'sweet potato'	'I grind'	'I eat'

(7) Five levels: Kam (Shidong) (China) (Edmondson and Gregerson 1992)

ʈa^{11}	ʈa^{22}	ʈa^{33}	ʈa^{44}	ʈa^{55}
'thorn'	'eggplant'	'father'	'step over'	'cut down'

The contrasts in (7) show that it is sometimes difficult to give names and use accent marks for each tone level, in which case it is more practical to indicate

pitch levels by numbers (5 = highest, 1 = lowest). In addition to such level tones, some languages also have contour tones which either rise, fall, or both. The best known such example is Standard Mandarin, whose four-way contrast is frequently exemplified by means of the minimal quadruplet in (8).

(8)				
	Tone I	(high level)	ma^{55}	'mother'
	Tone II	(rising)	ma^{35}	'hemp'
	Tone III	(falling-rising)	ma^{214}	'horse'
	Tone IV	(falling)	ma^{51}	'scold'

By exploiting other contours as well as differences in phonation, for example, breathiness or glottalization, a language can have even more tonal contrasts on monosyllables, as in the following Trique (Itunyoso) [Mexico] examples (Dicanio 2006):

(9)	Level		Falling		Rising	
	ββe^{4}	'hair'	li^{43}	'small'	yãh^{45}	'wax'
	nne^{3}	'plough (n.)'	nne^{32}	'water'	yah^{13}	'dust'
	nne^{2}	'to tell lie'	nne^{31}	'meat'		
	nne^{1}	'naked'				

As many people do not speak a native language with tone, tonal contrasts such as the above can be quite unfamiliar to linguists and language learners alike:

> Most language students, and even a shocking number of linguists, still seem to think of tone as a species of esoteric, inscrutable, and utterly unfortunate accretion characteristic of underprivileged languages – a sort of cancerous malignancy afflicting an otherwise normal linguistic organism. Since there is thought to be no cure – or even reliable diagnosis – for this regrettable malady, the usual treatment is to ignore it, in hope that it will go away of itself. (Welmers 1959: 1)

In this chapter we are concerned with whether and, if so, how tone is different from other aspects of phonology. While few phonologists would like to be identified with the above caricature, which Welmers (1973: 77) felt compelled to repeat 14 years later, there are occasional indications that a shocking number of linguists do indeed feel that tone is different from segmental phonology in rather dramatic ways. While we will conclude that there ARE important differences, let us first reject three rather extreme misconceptions about tone which are sometimes expressed:

(i) *Tone cannot be studied the same way as other phonological phenomena*. Upon encountering their first tonal experience, even seasoned field workers have asked me: "How can I tell how many tones my language has?" Each time this happens I am tempted to answer back with the rhetorical question: "How can you tell how many vowels your language has?" I have seen investigators try to discover the tonal categories by first eliciting long utterances, and then marking the relative

pitch changes between syllables, as one might initially do when approaching intonation. At best, this complicates the task. As in the case of voicing, nasality, vowel length, and other phonological contrasts, the normal technique is to first elicit individual words to determine the phonetic properties, and ultimately the phonemic contrasts. In the case of tone, this might then yield the tonal minimal pairs, triplets, quadruplets, and quintuplets seen in (4)–(8) above. Other languages may offer fewer minimal pairs or require specific contexts or *frames* in which the full range of contrasts can be discerned. For example, as seen in (10), the largely monosyllabic words of Hakha Lai [Burma] are pronounced with either HL falling (ˆ) or level L (ˋ) tone in isolation (Hyman and VanBik 2004):

(10)		in isolation		after ka=	‘my’
	a.	hmâa	‘wound’	ka hmâa	‘my wound’
		lûŋ	‘heart’	ka lûŋ	‘my heart’
	b.	kêe	‘leg’	ka kěe	‘my leg’
		kôoy	‘friend’	ka kŏoy	‘my friend’
	c.	sàa	‘animal’	ka sàa	‘my animal’
		kàl	‘kidney’	ka kàl	‘my kidney’

However, as seen in the forms on the right, the falling tone nouns split into two classes when following toneless proclitics such as ka= ‘my’: the nouns in (10a) remain HL, while those in (10b) are pronounced with an LH rising (ˇ) tone. The natural conclusion to draw is that Hakha Lai has three underlying tones, /HL, LH, L/, and an LH → HL rule that applies after pause (and in certain other environments – see Hyman and VanBik 2004).

Such observations do not make tone different from other aspects of phonology – one has but to think of the languages which merge segmental contrasts on words in isolation, for example, the voicing contrast on German *Rat* ‘advice’ vs. *Rad* ‘wheel’. In studying tone we need to be rigorous and comprehensive, just as we would have to be in attempting to analyze anything else. As we shall see, the issues that come up in the study of tone are quite complex, with more processes being available to tone, particularly at the phrase level, than to segmental phonology (cf. Section 4.2). However, since tones enter into paradigmatic contrasts very much like consonant and vowel features, we need not seek new methodologies. As Welmers (1959: 9) put it: “The more information we acquire about even the most complex tone systems, the more encouragement we receive that we already have the equipment needed to handle them.” In short, tone CAN be studied just like other aspects of phonology.

(ii) *Tone cannot mark certain things*. A second misconception is that tone is used only to mark certain things. Most of the examples in (4)–(10) illustrate the LEXICAL function of tone: Different monomorphemic nouns, verbs, and so on, differ only in tone. However, the Somali examples in (3) show that tone can also be implicated in grammar, that is, it can have a MORPHOLOGICAL function. Further examples of morphological tone are seen in (11)–(13).

(11) Tone marking person in Zapotec (Macuiltianguis) (Mexico) (Broadwell 2000: 7)

a.	1st person:	bè-xáttā?-yà?-nà	'I ironed it'
		bè-gállá?-yà?-nà	'I hung it up'
b.	3rd person:	bè-xàttā?-nà-nà	'he ironed it'
		bè-gàllá?-nà-nà	'he hung it up'

(12) Tone marking tense-aspect in Nambikuára (Brazil) (Kroeker 1977: 129)

a.	present:	xyău-nără	'he is staying'	săsŏ-nără	'he is taking'
b.	past:	xyàu-nără	'he stayed'	săsô-nără	'he took'

(13) Tone marking negation in Igbo (Aboh) (Nigeria) (elicited by the author)

a. /ò jè kọ̀/ → ò jè kọ̀ 'he is going'
b. /ó jè kọ̀/ → ó jé kọ̀ 'he isn't going' (H of subject /ó/ spreads onto /jè/)

In (11) we see that the first-person Zapotec verbs begin with an H tone, while the third-person verbs begin with L. In (12) the only difference between the present and past tense is tonal in Nambiquára (where ˆ and ˇ mark HL falling and LH rising tones, respectively). Finally, in the Aboh dialect of Igbo, the only difference between the affirmative and negative utterances in (13) is the tone on the third-person subject pronoun /o/ (whose H spreads onto /jè/ 'go' in the second example). It is thus clear that tone can have a grammatical function as well as a lexical one.

The question is whether there are grammatical notions that tone cannot mark. In his Presidential Address at the 2004 Linguistic Society of America Annual Meeting, Ray Jackendoff proposed in passing the following alleged universal: "No language uses tone to mark case." The Maasai (Kenya) data in (14), however, provide a rather clear counterexample to this claim (Tucker and Ole Mpaayei 1955: 177–184):

(14)

	nominative	accusative	
class I:	èlùkùnyá	èlúkúnyá	'head'
	èncùmàtá	èncúmátá	'horse'
class II:	èndéròní	èndèróní	'rat'
	ènkólòpà	ènkòlópà	'centipede'
class III:	òlmérégèsh	òlmérègèsh	'ram'
	òlósówùàn	òlósòwùàn	'buffalo'
class IV:	òmótònyî	òmótònyî	'bird'
	òsínkìrrî	òsínkìrrî	'fish'

In Maasai, native nouns usually consist of a gender prefix (masculine sg. *ol-*, feminine sg. *en-*) followed by a stem of one or more syllables. Except for class IV, nouns take different tones in nominative vs. accusative case. Although there are other complications, the four-syllable nouns which are cited show the following: in class I, the nominative has a single H on the last syllable, while the accusatives

have one L followed by all H syllables. In class II, nouns have a single H which is realized on the first stem syllable in the nominative vs. the second stem syllable in the accusative. In class III, the nominative has an H tone on the first and second stem syllables, while the accusative has an H only on the first stem syllable.

While Maasai is rather clear on the issue of tone marking case, the proposal is all the more surprising as so much of the tonal discussion in the 1960s and 1970s concerned the analysis of the so-called associative tone in certain West African languages such as Igbo (see Williamson 1986 and references cited therein). One analysis is that the associative marker consists of an H tonal morpheme which, as seen in the examples in (14), is assigned to the left in Central Igbo, but to the right in Aboh Igbo (Hyman and Schuh 1974: 98–99):

(15)	Central Igbo:	àgbà	+ ´ +	èŋwè	→	àgbá èŋwè	'jaw of monkey'
	Aboh Igbo:	ẹ̀gbà	+ ´ +	èŋwè	→	ẹ̀gbà éŋwè	'jaw of monkey'

In Igbo, the /L-L/ words *àgbà* ~ *ẹ̀gbà* 'jaw' and *èŋwè* 'monkey' are pronounced L-L in isolation. As seen, according to the dialect, an intervening H tonal morpheme is realized either on the preceding or following TBU in the associative construction. But what is this "associative morpheme" if not a genitive case marker? The more pressing question is why anyone would seek to limit the kinds of constructions or semantic notions that tone can mark. The failure here is to appreciate the full morphological nature of tone: If a tone can be a morpheme, then it can do anything that a morpheme can do. This follows from the fact that tonal morphemes most commonly derive historically from earlier segmental+tonal morphemes whose segments have been lost (cf. Section 3). Thus, anything that can be marked by a segmental+tonal morpheme, can also be marked by a tonal morpheme. In other words, "tonal morphology . . . exhibits essentially the same range of morphological properties as in all of segmental morphology" (Hyman and Leben 2000: 588).

The alternative view, that tone is better suited to express certain ideas rather than others, has, however, occasionally also been expressed:

> In a tone language, tone is not a purely harmonic or musical element, it is the expression of a thought, of an idea; it belongs to the intellectual domain, such that we can formulate the following axiom: SEEK THE IDEA AND YOU WILL HAVE THE TONE. (Stoll 1955: 5; my translation, his emphasis)

Along these lines, Stoll suggests that H tone indicates "everything that is woman, female, feminine" while L tone indicates "everything that is masculine, male, man" (p. 156). While certain languages occasionally exhibit indications of tonal iconicity (cf. Ratliff 1992), there is no reason for H vs. L tones to signify feminine and masculine any more than voiceless vs. voiced consonants (with which these tones often correlate, respectively). Once again, tone is not different from other phonological features.

(iii) *Tone is expendable*. The third misconception is that tone is somehow less essential in a language than other phonological features. This is seen in the way

tones and tone systems are sometimes dealt with in the literature. First, many descriptive and pedagogical grammars do not analyze or present the tone system in any detail. Those that do may have a section on tone, but then elect not to transcribe tone in the rest of the study. Different excuses are variously provided: (i) the tones are not important because they do not have a heavy functional load (i.e. there are few minimal pairs); (ii) the tones are a typographical inconvenience (e.g. an accent mark gets in the way of marking nasalized vowels with a tilde); (iii) the tones have not yet been analyzed; (iv) in order to learn the tones you have to listen to a native speaker anyway; (v) native speakers prefer not to write the tones, and so on. There are, of course, practical orthographies that do not mark tone, just as there are those which fail to mark all of the segmental contrasts. While there is a growing experimental literature on whether it is advantageous to readers if tone is marked, and if so, when and how (see, for example, Bird 1999a, b), the minimal contrasts cited in the above numbered examples establish that tone is extremely important in many, if not most, languages which have a tone system.

The view that tone is expendable is not limited to those designing practical orthographies. In general linguistic work, even phonologists commonly cite language data without indicating the tones. The most egregious cases occur when linguists REMOVE the tones from tone-marked examples cited from other sources, sometimes adding in a footnote that they have done so because "the tones are not relevant to the current study." Perhaps this is also encouraged by an awareness that stress, another prosodic property (which may be predictable or have a low functional load), is also often omitted from linguistic transcriptions, as it is from many practical orthographies. However, no other phonological feature is treated with such indifference as tone: contrastive voicing on consonants or contrastive rounding on vowels is never removed from original fully-marked examples because voicing/rounding "is not relevant to the current study." On the other hand, there are speech situations where the speakers themselves omit the segments in favor of the tones:

> . . . when, for some physical reason, it is inconvenient for a Nambikuára to separate his teeth, he may still participate in a conversation by talking through closed teeth using tone as his principle means of communication. (Kroeker 1977: 133–134)

The inherent importance of tone should therefore not be underestimated.

As I have tried to indicate, although the above three misconceptions are sometimes explicitly expressed, they are more frequently implicit in the way linguists go about doing their work. Whether overtly expressed or not, the evidence is that tone is thought of as something different from the rest of what phonologists – or linguists in general – study. These misconceptions are both extreme and wrong-headed. However, the question still remains. Is tone significantly different from other aspects of phonology? For example, is tone more independent from other phonological features than these latter are from each other? Is tone capable of greater variability than other features? Are there special processes which are found only in tone systems? Or, as Leben (1973b: 117) once put it:

> Is tone such a special phenomenon that it must be viewed as a feature on morphemes or larger units in some languages, as a feature on syllables in others, and as a feature on segments in still others? If so, then there is something left to explain: namely, why tone, unlike any other linguistic entity we know anything about, is capable of this many different types of representation.

Such questions are taken up in the following sections.

3 The Autosegmental Insight

Over the past several decades there has been an exponential increase in the documentation, analysis, and theoretical understanding of tone systems from throughout the world. Numerous formal proposals have been made concerning tone features and feature geometry, register effects such as downstep, and the interaction between tone and stress (de Lacy 2002), among other issues. In addition, the treatment of globality issues in tone has led to the development of Optimal Domains Theory (Cassimjee and Kisseberth 1998a). However, as mentioned in Section 1, the greatest impact of tone on phonological theory occurred in the 1970s when it provided the model for autosegmental phonology (Goldsmith 1976a, b). As a preliminary to the question of whether tone is "different," it is therefore appropriate to begin by considering the fundamental insight of autosegmental tonology, stated in (16).

(16) Tones are SEMI-AUTONOMOUS from the tone-bearing units on which they are realized.

This view constituted a reaction to the "segmental" approach of standard generative phonology (Chomsky and Halle 1968), which represented consonants and vowel segments in terms of a single matrix of binary distinctive features. For example, the vowel /a/ could be characterized by the vertical array of the binary feature values [–cons], [+back], [–round], and [+low], as in (17).

(17) a. High tone [á] b. Rising tone [ǎ]

$$\begin{bmatrix} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ +\text{HIGH} \end{bmatrix} \qquad \begin{bmatrix} -\text{cons} \\ +\text{back} \\ -\text{round} \\ +\text{low} \\ +\text{RISING} \end{bmatrix}$$

The question was how to represent the various level and contour tones in terms of features. An ad hoc response was to add tonal feature values to the segmental matrices in (17), for example, [+HIGH] for H tone, [+RISING] for rising tone (the capitals being used so as not to confuse tonal and vowel height features). However, there are two problems with the representations in (17).

The first problem stems from the inherent claim that tones are inseparable features on segments, whereas tones have considerable autonomy from their TBUs. As argued by Leben (1973a, b), some languages have a limited number of suprasegmental "tonal melodies" which must be abstracted away from the TBUs on which they are realized. Although his example was Mende (cf. Leben 1978), I cite examples in (18) which illustrate the corresponding five tonal melodies of closely related Kpelle (Liberia) (Welmers 1962: 86):

(18) a. High throughout

pá	'come'	ɓóa	'knife'
láa	'lie down'	píli	'jump'

b. Low throughout

kpòo	'padlock'	kpàki	'loom'
tɔ̀nɔ	'chisel'	tòloŋ	'dove'

c. High followed by low (low begins on the next vowel if there is one)

yɛ̂	'for you'	tôa	'pygmy antelope'
kpôŋ	'door'	kâli	'hoe'

d. Mid throughout

kpɔŋ	'help'	sua	'animal'
see	'sit down'	kali	'snake'

e. Mid with first vowel, then high followed by low

tɛɛ̂	'black duiker'	konâ	'mortar'
yuɔ̂	'axe'	kpanâŋ	'village'

In the above, I have adopted Welmers' practice of using only one tone mark per word (M is unmarked). He thus writes /kâli/ for what is pronounced [kálì] 'hoe', that is, H-L. Second, there is no difficulty reducing Kpelle to an underlying two-level system: The M that occurs in the MHL melody in (18e) is straightforwardly analyzed as an L which is raised before H (cf. Section 4.2), and the "mid throughout" melody in (18d) is underlyingly /LH/, as is seen when two "mid throughout" words occur in sequence:

> In mid-mid, for the dialect being described here, the first mid has a slightly rising allotone . . . In some areas, the first mid is level, but the second mid begins a little higher and drops quickly to the level of the first. In still other areas, both phenomena occur: the first mid ends a little higher, and the second begins a little higher. In all cases, the conjunction of two mids is accompanied by an upward pressure. (Welmers 1962: 87, Note 2)

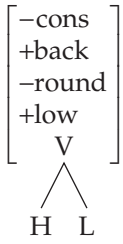
The important observation is that only five tone patterns (or "melodies") are possible independent of the number or nature of the TBUs: /H/, /L/, /HL/, /LH/, /LHL/. If, on the other hand, each TBU were capable of carrying an independent, underlying /H/, /L/, rising, or falling tone, we would expect $4 \times 4 = 16$ combinations on two TBUs, rather than the five that are observed. In this sense the tones are autonomous from the TBUs.

The second problem with (17) has to do with the representation of contour tones. Features such as [RISING] and [FALLING] encode a change within the segment which is not made explicit. In a two-level tone system, a rising tone typically acts as if it is a sequence of L+H realized on a single TBU, while a falling tone acts as if it is an H+L sequence on a single TBU. We see this in the following noun forms from Mende (Sierra Leone) (Leben 1978), which are arranged according to the same five tonal melodies as in Kpelle:

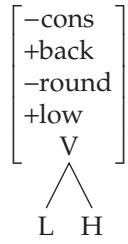
(19)			base noun		+ hu 'in'	+ ma 'on'
	a.	/H/	kɔ́	'war'	kɔ́-hú	kɔ́-má
	b.	/L/	bɛ̀lɛ̀	'trousers'	bɛ̀lɛ̀-hù	bɛ̀lɛ̀-mà
	c.	/HL/	mbû	'owl'	mbú-hù	mbú-mà
	d.	/LH/	mbǎ	'rice'	mbà-hú	mbà-má
	e.	/LHL/	nyàhâ	'woman'	nyàhá-hù	nyàhá-mà

As seen, the two locative enclitics *=hu* and *=ma* take their tone according to the tone of the preceding noun. When the noun ends in an H or L, as in (19a, b), the enclitic takes the same tone. When the noun ends in a falling (HL) tone, as in (19c, e) or a rising tone, as in (19d), the two parts of the contour "split": the first part goes on the last syllable of the noun, while the second goes on the enclitic. We thus see that when there is an extra available syllable, a falling tone maps as an H-L sequence and a rising tone maps as an L-H sequence. This then provides the evidence that contours should not be characterized by features such as [FALLING] and [RISING] but rather as sequenced level tone features, as in (20).

(20) a. Falling tone [â]

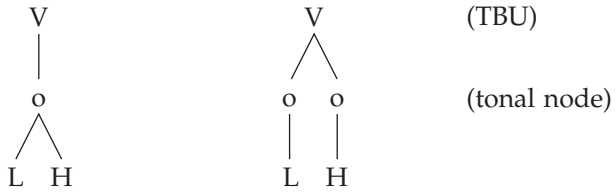


b. Rising tone [ǎ]



As proposed by Goldsmith (1976a, b), the H and L tones (or tone features) are represented on a separate tonal tier. Since they are both linked to the same TBU, the result is a contour tone. Other evidence that contours consist of two (occasionally more) independent tones linked to the same TBU can be cited from tone systems throughout the world. On the other hand, Yip (1989, 2002) argues that contour tones should be analyzed as units in certain Chinese dialects where the sequenced tone features appear to function as units. In such cases a rising tone would be represented roughly as in (21a), where the two tone features, L and H, are linked to a single tonal node.

(21) a. contour tone b. tone cluster



This is what Yip refers to as a true contour tone vs. the more common "tone cluster" representation in (21b), where each tone has its own tonal node.

With such representations, we now understand the meaning of the "semi-autonomy of tone": Tones are on a separate tier, but they are linked to their TBUs by association lines. Originally the proposal was that there were two tiers, a segmental tier and a tonal tier, but further elaborations were proposed to capture (i) the feature geometry of segments (Clements 1985; Clements and Hume 1995); (ii) the feature geometry of tones (Bao 1999; Snider 1999; Yip 1995, 2002), and (iii) the moraic and syllabic structures which organize the segments and serve as the TBUs to which the tones link. All of this was made possible by the autosegmental insight that tones are semi-autonomous from their TBUs. Support for this position has come from three general observations.

The first argument is that there is a NON-ISOMORPHISM between the two in the sense that the tones and TBUs do not necessarily synchronize: As we saw in (20) and (21), more than one tone can link to the same TBU, in which case we obtain a contour tone. The reverse, where one tone links to two TBUs, is also possible. As a result, Kukuya (Congo) (Paulian 1975; Hyman 1987) contrasts two kinds of H-H words. As seen in (22a), both *má-bá* 'they are oil palms' and *wátá* 'bell' are pronounced H-H in utterance-medial position:

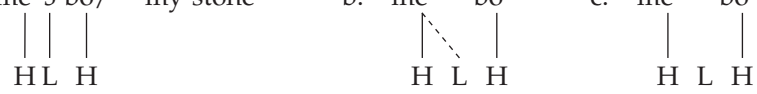
(22) a. Medial b. Prepausal

In (22b), however, the two words are realized differently before pause. As seen, there is an H → M rule which affects the last H feature before pause, not just the last H TBU. The contrasting representations in (22a), which had no equivalence in pre-autosegmental tonology, provide the structural difference that results in the surface opposition of H-M vs. M-M before pause. This difference largely correlates with a morphological difference: *má-bá* consists of two morphemes while *wátá* consists of one. In general, the double representation seen in *má-bá* is possible only when each H belongs to a different morpheme, while the branching representation of *wátá* is expected of single morphemes. There are exceptions in both directions, but in general the Obligatory Contour Principle (OCP) (Leben

1973a; Goldsmith 1976a) prohibits sequences of identical tonal features within the same morpheme.

The second argument for the semi-autonomy of tones from their TBUs concerns STABILITY EFFECTS. When a TBU is deleted, its tone is not necessarily deleted, but may either be relinked to another TBU, or it may "float" and have an effect on other tones. Both possibilities can be observed in Twi (Ghana) (Schachter and Fromkin 1968) in (23).

(23) a. /me ɔ-bo/ 'my stone' b. mê bó c. mé ↓bó



The input in (23a) consists of a /H-L-H/ sequence. When the historical L tone prefix /ɔ̀-/ is deleted, there are two options: In (23b), as indicated by the dotted association line, some speakers free-associate the L to the H tone pronoun /mé/ 'my' to form an HL falling tone. The more common option in (23c), however, is for the L to stay afloat and cause a lowering or "phonemic" DOWNSTEP of the following H of the root /bó/ 'stone'. Since the deletion of a vowel does not require the deletion of its tone, we have a strong validation of the decision to represent the tone on its own tier.

The derivation of a contour tone or downstep from the loss of a TBU is very common. Another example of the latter comes from Bamileke-Dschang (Cameroon) (Tadadjeu 1974):

(24) a. séŋ è séŋ 'the bird of the bird' → séŋ ↓séŋ
H L H H L H

b. séŋ ↓séŋ ↓séŋ ↓séŋ ↓séŋ 'the bird of the bird of the bird of the bird . . .'
H L H L H L H L H

In very deliberate speech, the input associative (genitive) marker /è/ 'of' is pronounced as an L tone [è]. When the vowel is deleted, however, the stable L tone produces a downstep on the possessor. The artificial, but grammatical, sequence in (24b) shows that the downstep effect is iterative: each floating L tone conditions a successive drop on the following H.

The Bamileke-Dschang example leads naturally into the third argument for the semi-autonomy of tones from their TBUs, the possibility of ZERO REPRESENTATION: a morpheme can consist solely of a tone without a TBU and without segmental features. Tonal morphemes of course derive from full syllables which have deleted. At a point where future Bamileke-Dschang speakers can no longer pronounce the /è/ in (24a), they will have derived an L associative tonal morpheme, much like the H associative tonal morpheme that was seen in Igbo in (15).

While a tone can be a morpheme until itself, other floating tones can be lexical. Such a situation obtains in Peñoles Mixtec (Mexico), whose TBUs show an underlying contrast between /H/, /L/, and /Ø/ (Daly and Hyman 2007):

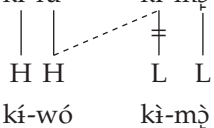
(25) a. /kɨtɨ/ ‘animal’
/nǰuši`/ ‘chicken’

b. kɨtɨ ditó ‘uncle’s animal’
nǰuši di̱tó ‘uncle’s chicken’

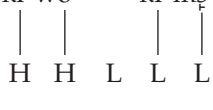
Both of the nouns in (25a) are toneless, pronounced identically with a low falling contour before pause or an L tone. However, as seen in (25b), the realization of /ditó/ ‘uncle’ is different after the two nouns. In the case of ‘uncle’s animal’, the three toneless TBUs are pronounced on a level mid pitch. In the case of ‘uncle’s chicken’, the floating L of ‘chicken’ links to the first syllable of ‘uncle’ and the two toneless TBUs of *nǰuši* continue to be pronounced with a low falling contour. As seen in the underlying forms in (25a), nouns such as */nǰuši`/* ‘chicken’ have a floating L tone after them (which derives from the loss of Proto-Mixtec final glottal stop, see Longacre 1957; Dürr 1987).

While the effects of lexical floating tones can be discerned in tone systems from all parts of the world, they are particularly common in Mexico and West Africa. Besides Peñoles, other Mixtec languages show the effects of a lexical floating L, for example, Atatláhuca (Mak 1953); a lexical floating H, for example, Chalcatongo (Hinton *et al.* 1991) and San Miguel el Grande (Pike 1948, Goldsmith 1990: 20–27); or both, for example, Acatlán (Pike and Wistrand 1974), Jacaltepec (Bradley 1970), and Magdalena Peñasco (Hollenbach 2004). Within Africa, the Grassfields Bantu languages are particularly well known for their complex floating tone systems (Voorhoeve 1971; Hyman and Tadadjeu 1976). In Aghem (Cameroon) (Hyman 1979b), although the two nouns *kɨ́-fú* ‘rat’ and *kɨ́-wó* ‘hand’ are pronounced identically as H-H in isolation, they have different effects on the tone that follows:

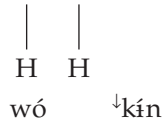
(26) a. kɨ́-fú kɨ̀-m̩̀ ‘one rat’



kɨ́-wó kɨ̀-m̩̀ ‘one hand’



b. fú kɨ́n ‘this rat’



wó ↓kɨ́n ‘this hand’

As indicated by the dotted line in (26a), the H tone of the root *-fú* spreads onto the prefix of the numeral ‘one’ (pronounced *kɨ̀-m̩̀* in isolation), whose L tone then delinks. The H tone of the root *-wó* ‘hand’, however, does not spread. This is because it is followed by a floating L which belonged to a historically lost second syllable (cf. Proto-Bantu **-bókò* ‘hand’). While much of the older work on Mixtec languages would have simply divided up H-H nouns into an arbitrary class A vs. class B, the floating L analysis has the advantage that it naturally accounts for the additional difference observed in (26b). Here the demonstrative *kɨ́n* ‘this’ (which conditions the deletion of the noun class prefix *kɨ́-*) is realized H after *fú* ‘rat’, but as a downstepped H after *wó* ‘hand’. As we saw in (23) and (24), a floating L frequently conditions downstep, as it does in Aghem. The floating L of /kɨ́-wó`/ is thus not circular. It blocks H tone spreading onto a following L and

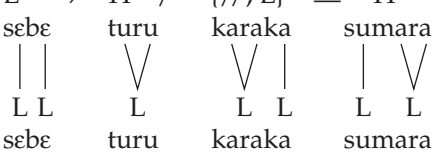
conditions downstep on a following H, something which would not be naturally captured by a class A/B diacritic account.

It should be clear from the foregoing that the autosegmental representations, which express the traditional intuition of the semi-autonomy of tone, provide a more explanatory analysis in many cases. This does not mean that all tonal phenomena will have the properties illustrated in the preceding paragraphs. In some languages, when a vowel is deleted, its tone shows a stability effect, as above, while in others, for example, Shilluk (Gilley 1992: 164), the tone is deleted with the vowel. In addition, some languages may not distinguish intra- vs. heteromorphemic representations as in (22a). Worse yet, they may violate the OCP and show contrasts WITHIN morphemes. Odden (1982, 1986), for example, argues that in Shambala *nyóká* 'snake' has one doubly linked H tone vs. *ngó↓tó* 'sheep', which, in violation of the OCP, has two underlying H tone features, the first causing the second to downstep. Similar issues arise concerning underlying L tone in Dioula (Odienné) (Ivory Coast) (Braconnier 1982):

(27)	a. before pause	b. before H	
	sɛ̀bɛ̀	sɛ̀bɛ́	'paper'
	tùrù	túrú	'oil'
	kàràkà	kàràká	'bed'
	sùmàrà	sùmárá	'soumbala' (a spice)

As seen in (27a), the four monomorphemic nouns are pronounced all L before pause. When followed by an H, as in (27b), both the bisyllabic and trisyllabic nouns show two patterns: either one or two L tone syllables become H. The analysis may depend on how the rule is formulated. If the rule is as in (28a), the observed differences can be represented in one of two ways:

(28) a. L → H / {//, L} __ H

b.


c.

In (28b) a noun may have one vs. two L tone features, where only the L feature which precedes the H is raised to H. Depending on one's theoretical assumptions, this analysis potentially has two problems: (i) the forms with two Ls violate the OCP intramorphemically; (ii) the rule in (28a) is formulated as a feature-changing rule rather than a tone spreading rule – which is the most common way to express tonal assimilations (Section 4.1). In response to the first problem we might instead propose the representational differences in (28c). As seen, there is now only one /L/, linked either to one or two TBUs. As also seen, this /L/ can be preceded

by one or more toneless TBUs. The idea here is that the raising rule affects only those TBUs that are prelinked to the /L/.

Proposals of underlying /H, L, Ø/ systems go back at least to Pulleyblank's (1986) treatment of Margi (Nigeria), which realizes /Ø/ as [H] or [L] and Yoruba (Nigeria), which realizes /Ø/ as [M] (cf. Akinlabi 1985). If /L/ contrasts with /Ø/ in Dioula, still another alternative is to propose the underlying representations as in (29a), where the nouns end in one or more toneless syllable:

(29) a. sɛbɛ turu karaka sumara

L L L

b. {//, L} $\underline{\sigma}^n$ # σ

H

With these representations, the tonal assimilation rule can be reformulated as an anticipatory spreading rule, as in (29b), where the underlined $\underline{\sigma}^n$ represents one or more toneless TBUs. Which one of the above is the most satisfactory account of the facts can be determined only by an in-depth analysis of Dioula d'Odienné, which is not an isolated case: We face similar analytic choices in Acatlán Mixtec (Mexico) (Pike and Wistrand 1974), where some L-L words become H-L, others H-H after what is most naturally analyzed as a floating H tone. Such representational issues pervade tonological analysis, perhaps even more than in the analysis of segmental systems.

4 Is Tone Different?

With the semi-autonomy of tone now firmly established, we turn to the question of whether tone is different from the rest of phonology. In her textbook on tone, Yip (2002: 65) observes the following possible differences between tone and segmental phonology:

(30) "Tone differs from many other phonological features in the following ways, rarely or never observed in more familiar consonant or vocalic features:

a. Mobility: movement away from point of origin
b. Stability: survival after loss of original host segment
c. One-to-many: a single tonal feature shared by two or more segments
d. Many-to-one: multiple tonal features surfacing on a single host segment
e. Toneless segments: potentially tone-bearing segments that never acquire phonological tone"

As seen, Yip's summary essentially recapitulates the autosegmental nature of tone, such that tone would seem to have more semi-autonomy than consonant or vowel

features. We might therefore say that *tone is like segmental phonology in every way – only more so!* The phrase "only more so" can have two meanings: (i) QUANTITATIVELY more so: tone does certain things more frequently, to a greater extent, or more obviously (i.e. in a more straightforward fashion) than segmental phonology. (ii) QUALITATIVELY more so: tone can do everything segments and non-tonal prosodies can do, but segments and non-tonal prosodies cannot do everything tone can do. In fact, both of these characterizations are correct, as we shall see in the following two subsections.

4.1 Quantitative Differences Between Tonal and Non-tonal Phonology

In this section I discuss the following properties of tones, as they apply within the word domain: spreading, local shifts, non-local shifts, plateauing and polarity. Each of these is frequently attested in tone systems, but only the first is robustly attested in segmental phonology. I begin with the most common tonal process, tone spreading, whose properties are summarized in (31).

(31) Horizontal assimilation (or tone spreading) (Hyman 1975: 223)

a. Natural	b. Unnatural
L-H → L-LH	L-H → LH-H
H-L → H-HL	H-L → HL-L
(perseverative)	(anticipatory)

As seen, the term "horizontal assimilation" refers to cases where a full tone spreads onto a neighboring TBU. (Register effects or "vertical assimilations" are discussed in Section 4.2.) As is well known, even when there is no tone spreading, tonal targets tend to be reached late within their TBU: "Late realization of tonal targets has been demonstrated both for languages in which tones are lexical . . . and for those in which they are intonational . . ." (Kingston 2003: 86). As a consequence, phonological tone spreading tends to be perseverative, and where anticipatory, spreading is much rarer and has a quite different character (Hyman 2007b). This stands in marked contrast to what is usually said about segmental assimilations:

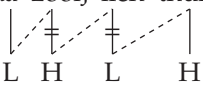
> In regular conditioned sound changes, the conditioning factor is far more frequently a sound which follows than one which precedes. (Greenberg 1957: 90)

> I examined 365 segmental assimilatory rules culled from 60 languages . . . documented in the Stanford Phonology Archive. 195 of these rules involved anticipatory assimilation of a segment to a following segment. 89 of these involved the perseverative assimilation of a segment to a preceding segment . . . The conclusion must be that segmental assimilation is generally anticipatory . . . (Javkin 1979: 75–76)

As seen in (32a), alternating sequences of input Hs and Ls undergo both H tone-spreading (HTS) and L tone spreading (LTS) in Yoruba (Laniran and Clements 2003: 207):

(32) a. /máyọ̀ mírà wé/ [máyo̰ mĭ râ wĕ] 'Mayomi bought books'

b. /kà zóoŋ lìen thúm/ [kà zòoŋ líen thŭm] 'my three big monkeys'



Since phrase-internal contour tones are permitted in Yoruba, the result is a sequence of falling and rising tones. This contrasts with the situation in (32b) from Kuki-Thaadow (NE India, Burma) (Hyman 2010). Since this language does not allow phrase-internal contours, LTS and HTS condition delinking of the original tone except on the final syllable. The result is a bounded tone SHIFT: both the first H and the second L in the input in (32b) are realized only on the following TBU.

A similar relation between tone spreading and shifting is seen in the closely related Nguni Bantu languages in (33), where the processes are unbounded:

(33)	a. Ndebele	b. Zulu	
	ú-kú-hlek-a	u-kú-hlek-a	'to laugh'
	ú-kú-hlék-is-a	u-ku-hlék-is-a	'to amuse (make laugh)'
	ú-kú-hlék-ís-an-a	u-ku-hlek-ís-an-a	'to amuse each other'

In (33a) the underlying (underlined) H tone of the initial prefix /ú-/ spreads up to the antepenultimate in Ndebele (Zimbabwe) (Sibanda 2004). The result is an H tone sequence spanning several syllables. However, as seen in (33b), the same H tone SHIFTS to the antepenultimate syllable in Zulu (South Africa) (Downing 1990: 265). In this case unbounded spreading + delinking has produced unbounded tone shift. Ndebele, thus, represents the older situation.

We can assume that both bounded and unbounded spreading have analogues in segmental phonology, where the most natural comparison is with processes such as vowel harmony. However, while it is very common for an underlying tone to shift several syllables to another position within the word (or onto a subsequent word, as will be seen in Section 4.2), there are very few cases reported where a segmental feature has this property. One such case in progress comes from Makonde (Mozambique) (Liphola 1999, 2001). As seen in (34), a process of vowel height harmony converts the applicative suffix /-il-/ to [-el-] after the mid root vowels /e/ and /o/:

(34)	underlying	VH-harmony	V-lengthening	V-reduction	
a.	/ku-pe̲t-il-a/	→ ku-pet-el-a	→ ku-pet-eel-a	~ ku-pa̲teela	'to separate for'
b.	/ku-po̲t-il-a/	→ ku-pot-el-a	→ ku-pot-eel-a	~ ku-pa̲teela	'to twist for'

While the height harmony process is widespread in Bantu, as is phrase-penultimate vowel lengthening, Makonde appears unique in allowing /e/ and /o/ to reduce to [a] in pre-penultimate position. As seen, both inputs are potentially realized

as *ku-pat-eel-a*, which suggests a local shift of the mid vowel height feature to the right. The additional examples in (35) show that the shift is potentially unbounded:

(35)	a. 'to not reach a full size for'	b. 'to cough for'	
	/ku-pelivilil-il-a/	/ku-kolumul-il-a/	
	ku-pelevelel-eel-a	ku-kolomol-eel-a	(no vowel reduction)
	ku-palevelel-eel-a	ku-kalomol-eel-a	(one application)
	ku-palavelel-eel-a	ku-kalamol-eel-a	(two applications)
	ku-palavalel-eel-a	ku-kalamal-eel-a	(three applications)
	ku-palavalal-eel-a		(four applications)

As seen, mid-vowel reduction to [a] follows vowel harmony and applies optionally to any number of mid vowels that precede the penultimate. The major constraint is that if an [e] or [o] is not reduced, it is not possible for another mid vowel to its right to be reduced (**ku-pelavelel-eel-a*, **ku-kolomal-eel-a*). This suggests either that reduction applies left-to-right or that there is a no-gapping constraint against the mid height feature (Archangeli and Pulleyblank 1994). As seen, the last variants of (35a, b) have the same property as the H tone shift in Zulu: In *ku-palavalal-eel-a* and *ku-kalamal-eel-a*, the mid feature originates in the underlined root vowel, but shifts several syllables to the penultimate. Although such segmental shifts are extremely rare (I do not know of another such case), Makonde shows that it is possible. Hence, the difference between tone, which frequently shifts, and segmental features, which rarely shift, is a quantitative one in this case.

Turning to anticipatory processes, we first note that bounded right-to-left tone spreading is extremely rare. As schematized in (36a), Inkelas and Zec (1988: 230–231) analyze Belgrade Serbian with a rule that spreads an H onto a pretonic mora, illustrated in the examples in (36b).

(36) a.

b.

/papríka/	→	pápríka	'pepper'
/raazlíka/	→	raázlíka	'difference'
/ne-ráadnik/	→	né-ráadnik	'non-worker'

Within Bantu, anticipatory local shift is also rare, but does occur in Kinande (Mutaka 1994):

(37)	a. /e-ri-túm-a/	→	e-rí-tum-a . . .	'to send'
	b. /e-ri-na-túm-a/	→	e-ri-ná-tum-a . . .	'to send indeed'

It should be noted, however, that anticipatory spreading and shifting are quite different from their perseverative counterparts (Philippson 1991: 180; Hyman 2007b: 18–28). Whereas the latter were said to derive from the phonetic tendency for tonal targets to be realized late, there is no corresponding phonetic tendency to realize tonal targets early. Instead, the above examples appear to have the property of anticipating prominent tones, in particular a /H/ tone which is opposed to /Ø/. Except as a phrasal property (Section 4.2), unbounded tone spreading is also rare,

and unbounded anticipatory tone shifting even more so. In some cases tonal anticipation is restricted to applying from a weak final to a strong penultimate syllable, for example, Chichewa /pez-á/ → *peézá* ~ *peéza* 'find!' (Kanerva 1989). This, then, suggests two different motivations for tones to spread and shift: the phonetic perseverative tendency and the attraction of a tone to a metrically strong position. We should thus expect more unbounded perseverative (vs. anticipatory) spreading/shifting to a metrical syllable, since such processes are doubly motivated.

If segmental features show more favor to anticipatory assimilation than tone, we should expect more anticipatory spreading and shifting. The numerous harmonies known as umlaut or metaphony fall into this category, as in the case of Servigliano Italian analyzed by Walker (2005: 918):

(38) a. verd-ó 'very green (m.sg.) vi̱rd-ú 'very green (m.pl.)'
b. kommonek-á 'to communicate' ku̱mmu̱ni̱k-ímo 'we communicate'

As seen, the mid vowels of the roots seen in (38a) assimilate in height to the high vowel of the following suffix in (38b). While most anticipatory cases involve suffix triggers, Esimbi (Cameroon) has a rather curious vowel height shift from root to prefix (Stallcup 1980; Hyman 1988), exemplified in (39).

(39) a. /u-ri/ → u-ri 'to eat' /u-mu/ → u-mu 'to drink'
b. /u-se/ → o-si 'to laugh' /u-kə̌/ → o-kǐ 'to beg' /u-mo/ → o-mu 'to go up'
c. /u-yɛ/ → ɔ-yi 'to wear' /ú-tă/ → ɔ́-tǐ 'to leave' /ú-mɔ/ → ɔ́-mu 'to sit'

The verbs in (39) consist of a verb root preceded by the infinitive prefix /u-/, which is specified only for rounding. As seen, the height feature of the root transfers to the prefix. This produces the minimal triplet involving the phonetic root [mu]. Stallcup (1980) hypothesizes that the prefix became accented, thereby driving an anticipatory vowel height harmony. Subsequently, root vowels reduced to [+high], which, if unmarked for height, can be characterized as delinking. Given the relatively small number of cases, it is not clear if anticipatory shifting favors segmental vs. tonal features.

Another process which is distinctly tonal is H tone plateauing. A number of tone systems prohibit *H-L-H or *H-Ø-H sequences (Cahill 2007), which Yip (2002: 137) refers to as *Trough. A common repair is for the non-H TBU(s) to be raised, such that an H tone plateau is created. In some languages, the process is limited to a single L TBU wedged between Hs, for example, Kihunde (Goldsmith 1990: 36) and Mamaindé [Brazil], about which Eberhard (2007: 297) writes "The heart of the tone sandhi issue in Mamaindé verbs resolves around sequences of HLH. There seems to be some sort of restriction against any HLH sequences in certain contexts (across the verb-stem/affix boundary). When this illegal sequence occurs, the intermediate L is always delinked." In other languages, for example, Amahuaca (Peru) (Russell and Russell 1959: 152) and Luganda (Stevick 1969; Hyman, Katamba, and Walusimbi 1987) multiple TBUs may undergo plateauing. The Luganda example in (40a) establishes that there is an H to L pitch drop on the last two syllables when the subject prefix is toneless /a-/ '3rd sg. (class 1)':

(40) a. a-a-tu-gul-ir-a → a-a-tu-gul-ír-à 'the one who buys it for us'
H L

b. a-ba-tu-gul-ir-a → a-bá-tú-gúl-ír-à 'the ones who buy it for us'
H H L H H

In (40b), where the subject prefix is H tone /bá-/ '3rd pl. (class 2)', the result is an H plateau of four TBUs (cf. also Section 4.2). The crucial point is that the two H tones can "see each other" at some distance. In fact, in some cases the plateauing process between Hs requires the deletion of one or more L features, which contrast with Ø (cf. (52)).

While such unbounded plateauing effects are quite common in tone, the question is whether comparable phenomena exist within segmental phonology. I know of only one example, mid vowel height plateauing in Yaka (Bantu; Democratic Republic of the Congo), illustrated in (41), based on van den Eynde (1968):

(41)		root + a	root gloss	applicative	causative	perfective	URs
	a.	kik-a	'obstruct'	kik-il-a	kik-is-a	kik-idi	/kik-ile/
		kul-a	'chase s.o.'	kud-il-a	kud-is-a	kud-idi	/kud-ile/
		kas-a	'bind'	kas-il-a	kas-is-a	kas-idi	/kas-ile/
	b.	keb-a	'pay attention to'	keb-il-a	keb-is-a	keb-ele	/keb-ile/
		sol-a	'clear bush'	sod-il-a	sol-is-a	sol-ele	/sol-ile/

In (41a) we see that the applicative, causative, and perfective suffixes all have the high vowel [i] when the root vowel is /i/, /u/, or /a/. In (41b), the causative and applicative are seen still to be *-il-* and *-is-* after the mid root vowels /e/ and /o/. Yaka would therefore appear to be lacking the widespread Bantu perseverative vowel height harmony process exemplified in Makonde in (34). However, the perfective suffix does show an alternation: *-idi* after /i, u, a/ vs. *-ele* after /e, o/. (The consonant alternation is due to a process whereby *l* → *d* before [i].) Why, then, should the harmony process be restricted to the perfective suffix, which often escapes height harmony in other Bantu languages? A number of arguments are presented in Hyman (1998) that the correct underlying representation of this suffix is /-ile/, as it is pronounced in many other Bantu languages. The harmony process thus targets high vowels which occur between two mid vowels. The forms in (42) show that mid vowel height plateauing applies to any number of high vowels which occur between mid vowels:

(42)		root + iC + a	stem gloss	perfective	URs
	a.	yed-ik-a	'taste'	yel-ek-ele	/yel-ik-ile/
		kos-ik-a	'add'	kos-ek-ele	/kos-ik-ile/
		yek-uk-a	'be separated'	yek-ok-ele	/yek-uk-ile/
		tob-uk-a	'be pierced'	tob-ok-ele	/tob-uk-ile/

b.	bet-idik-a	'lower'	bet-elek-ele	/bet-ilik-ile/
	kel-umuk-a	'turn around'	kel̲-omok-ele̲	/kel-umuk-ile/
			M -------- M	

Mid height plateauing represents a response to the precariousness of word-final /e/, which must either spread to another post-root vowel or be peripheralized to [i]. Thus, unlike most other Bantu languages, Yaka does not allow bisyllabic noun stems of the form CVCe (where V = any vowel). While H tone plateauing is quite frequent, the Yaka process is quite unique. It does, however, show that segmental plateauing is possible, however rare.

The last process to be considered in this subsection is polarity. Although having a number of manifestations, the best known case of tonal polarity occurs when an affix or clitic takes the opposite tone of its base or host. An example of this occurs in Margi (Nigeria) (Pulleyblank 1986: 203):

(43) a. hɔ̀gyì gṵ 'you are a Higi'
b. màrgyí gṳ̀ 'you are a Margi'

As seen, the subject clitic /gṳ/ 'you sg.' takes the opposite or polar value of the tone that precedes it: H after an L tone vs. L after an H. Note that in such cases, which are quite frequent, despite different possible solutions, it is often difficult to assign a unique underlying tone to the polar morphemes. The same point applies to cases of polar boundary tones: /H/ and /L/ are realized as LH and HL utterance-initially in Ticuna (Colombia) (Montes-Rodrigues 1995) and as HL and LH utterance-finally in Thlantlang Lai (Burma) (Hyman 2007a: 14). In these languages the boundary tone is polar to the adjacent lexical tone.

Since it would be arbitrary to propose a specific underlying tone in the above cases, tonal polarity differs from tonal (or segmental) dissimilation, where one of two identical specified features dissimilates (Hyman and Schuh 1974: 100). Thus, the unusual and mysterious low vowel dissimilation /CaCa/ → CeCa in Woleian and Marshallese and /CaCa/ → CiCa in Ere (Blust 1996) do not seem to be parallel. However, it is not out of the question that certain apparent polar effects owe their existence to a historical dissimilatory process. Consider, for example, the following alternating H tone pattern in Kirundi (Goldsmith and Sabimana 1986):

(44)	-sab- 'ask for'	-báz- 'ask (question)'	
a.	ku-sab-a	ku-báz-a	'to ask' (infinitive)
	ku-bi-sab-a	ku-bí-baz-a	-bi- 'them'
b.	ku-bí-mu-sáb-a	ku-bí-mu-báz-a	-mu- '(to) him'
	ku-bí-mu-kú-sab-ir-a	ku-bí-mu-kú-bar-iz-a	-ku- '(for) you'
	ku-há-bi-mú-ku-sáb-ir-a	ku-há-bi-mú-ku-bár-iz-a	-ha- 'there'
	'to ask him (for) them for you there'		

Ignoring the infinitive prefix *ku-* we see that up to three H tones may appear from the tone span that includes the object prefixes and the verb root. While Goldsmith and Sabimana account for the alternating H-L pattern is in metrical terms, another

way to conceptualize the pattern is to assume an earlier H tone plateau (as Furere and Rialland 1983 report for closely related Kinyarwanda), to which a left-to-right H-H → H-Ø dissimilatory process subsequently applies. Either way, what makes the Kirundi facts interesting are the forms in the left column of (44a). When a toneless verb root occurs with one or no object prefix, the whole infinitive is toneless (vs. the corresponding forms involving an H verb root). However, as soon as a second object prefix is added in (44b), we obtain not one H, but two. It would seem that as soon as one H is introduced by the morphology, it must be alternated within the tone span. While the Kirundi facts are reminiscent of alternating stress, I am unaware of any parallel case where a segmental feature is assigned on an alternating basis.

In this section we have examined several tonal phenomena as they apply roughly within the word domain. In each case it was suggested that the tonal process is more natural and frequent than its segmental analogue. One way to look at this is to say that tone is less restricted than segmental phonology. This interpretation receives support from a common restriction in segmental phonology which seems rarely to apply to tone: root control (Clements 1981). Whereas many harmony processes involve the assimilation of (underspecified) affixes to segmental features of the root, even word-level tonal assimilations seem largely to apply across-the-board. To take just one example, it has been oft noted that prefixes rarely, if ever, condition vowel harmony on a following root (Hall and Hall 1980: 227n). However, it is quite commonplace for the tone of a prefix to spread onto the following root, as seen in the Aghem examples in (45):

(45) a. HTS: /kí-kɔ̀ɔ̱̀ɔ̱/ → kí-kɔ̂ɔ̱̂ɔ̱ ‘cutlass’
b. LTS: /kì-téɛ/ → kì-tɛ̌ɛ ‘cricket’

It would appear that the perseverative tendency for tonal targets to be realized late overrides any counteracting tendency for prefixal tones to assimilate to roots. Since segmental assimilations were said to have an apparent anticipatory bias, we make the following statistical prediction: Roots will tend to assimilate to the tones of prefixes, but to the segmental features of suffixes. While we should not expect this prediction to be without exception, I believe that the asymmetry is quite real (Hyman 2002).

4.2 Qualitative Differences Between Tonal and Non-tonal Phonology

In the preceding section, the comparison of tonal to segmental processes was largely limited to the word domain. In this section we will consider tonal properties that apply across words. As will be seen, once we do so, the differences become more pronounced. We begin by considering “register effects” and then move on to discuss long-distance processes that apply at the phrase level.

In Section 4.1 it was said that tones tend to perseverate in “horizontal” assimilations. This is true only if we are looking at full tone assimilation. A second

possibility is for tones to undergo "vertical assimilation" or register adjustments. In this case, the asymmetries are quite different, as summarized below:

(46)

a. Compression

Input:	Anticipatory	Perseverative
L-H	M-H	L-M
H-L	? M-L	? H-M

b. Expansion

Input:	Anticipatory	Perseverative
L-H	? $^{\downarrow}$L-H	? L-$^{\uparrow}$H
H-L	$^{\uparrow}$H-L	H-$^{\downarrow}$L

As seen in (46a), either tone of a /L-H/ input can undergo a pitch register adjustment and ultimately become a third level tone: the L may be raised to M, the H may be lowered to M, or as we saw in Kpelle in (18d), both may apply in which case /L-H/ is realized [M-M]. The result is tonal COMPRESSION: the difference in the interval between the two output tones becomes smaller than in the input. An input /H-L/ does not show this effect. As indicated by the question marks, it is quite unusual for the H to be lowered or the L to be raised. In fact, as seen in (46b), the opposite effect of EXPANSION is observed: /H-L/ (but not /L-H/) may undergo raising of the H or lowering of the L. The raising process is seen in the Engenni (Nigeria) example in (47a) (Thomas 1974: 12):

(47) a. /únwónì/ 'mouth' → [únwő́nì]

b. /únwónì + ólíló/ 'mouth of a bottle' → [únwő́n ólíló]

In Thomas' analysis a single H TBU is raised to a "top" or superhigh (˝) tone before an L. In (47b) the L tone vowel /ì/ is elided, but still has the raising effect on the preceding tone. The result is a surface contrast between the H and superhigh tone.

H tone raising before L is a quite widespread phenomenon, also occurring for example in Gurma (Burkina Fasso) (Rialland 1981, 1983), Kirimi (Cahi) (Tanzania) (Hyman 1993), Edopi (Indonesia (Papua)) (Kim 1996), and Chinantec (Lealao) (Mexico) (Rupp 1990). Tesfaye and Wedekind (1990: 360) report that in Shinasha (Ethiopia) an H-L drop is realized "with about four semitones" while an L-H rise "is realized as a pitch increase of only two, sometimes three semitones." The question is whether there is a physiological reason for such differences or whether H-raising is a strategy for maximizing the tonal space – or both? Many languages have "downdrift" or "non-phonemic" downstep whereby the second H of an H-L-H sequence is realized on a lower pitch than the first. If there are enough transitions from H to L to H, Hs which are late in the utterance may become quite low. Since

it increases the tonal space, raising an H before an L could thus be useful as a pre-planning counterforce to processes which lower tones (Rialland 2001).

Some support for this interpretation is obtained from languages which raise H tone in anticipation of a contrastive downstep ($^{\downarrow}$H). In languages which contrast H-H vs. H-$^{\downarrow}$H there can be multiple downsteps within an utterance, with each $^{\downarrow}$H being produced at a lower pitch level than the previous H. Starting at a higher level could thus be quite useful. In this context consider the following pairs of examples from Amo (Nigeria) (Hyman 1979a: 25) and Luganda (personal notes with Francis Katamba):

(48) Non-local H → $^{\uparrow}$H in anticipation of (long-distant) phonemic downstep

a.	kìté úkɔ́ɔ́mí fínáwà	'the place of the bed of the animal'
	kì$^{\uparrow}$té úkɔ́ɔ́mí fíká$^{\downarrow}$lé	'the place of the bed of the monkey'
b.	à-bá-síb-á kígùùndú	'the ones who tie up Kigundu'
	tè-$^{\uparrow}$bá-síb-á $^{\downarrow}$kígùùndú	'they do not tie up Kigundu'

In each pair the second utterance contains a $^{\downarrow}$H which is lacking in the first. As indicated, but not usually transcribed, a sequence of Hs is quite audibly raised in anticipation of the downstep. Thus, between the first and second syllables, there is a step up of [+1] in the examples lacking a $^{\downarrow}$H vs. [+2] in the examples having a later $^{\downarrow}$H. While the raising of H before L appears to be local, it is striking how early the first H is raised in anticipation of the non-local H-$^{\downarrow}$H interval. As we shall see below, other such long-distance interactions are quite typical of phrasal tonology.

The properties of what I refer to as canonical downstep (Clements 1979; Hyman 1979a) are as follows: (i) H contrasts with $^{\downarrow}$H only after another ($^{\downarrow}$)H; (ii) $^{\downarrow}$H establishes a "ceiling effect" until the register is re-set: thus, a following H will be pronounced at the same pitch level rather than higher; (iii) there is a theoretically unlimited number of downstep pitch levels (H-$^{\downarrow}$H-$^{\downarrow}$H-$^{\downarrow}$H-$^{\downarrow}$H . . .). While such systems are best known from Sub-Saharan Africa, they are in fact found in languages from throughout the world, for example, Kuki-Thaadow (NE India, Burma) (Hyman 2010), Kairi (Rumu) (Papua New Guinea) (Newman and Petterson 1990), Mixtec (Coatzospan) (Mexico) (Pike and Small 1974), Tatuyo (Colombia) (Gomez-Imbert 1980), among many others. Downstep is, thus, a very natural tonal phenomenon. The question, then, is what, if anything, corresponds to downstep in segmental phonology? The brief answer: nothing. Again, we might look to vowel height for a parallel. Recall from (39) the transfer of vowel height features from root to prefix in Esimbi. The eight underlying root vowels are exemplified in (49) and as they are realized in their singular and plural forms:

(49)		URs	gloss	class 3 sg. /u-/	class 6 pl. /a-/
	a.	/-tili/	'end'	u-tili	o-tili
		/-wúsu/	'fire'	u-wúsu	o-wúsu
	b.	/-yembe/	'song'	o-yimbi	ɛ-yimbi
		/-góro/	'foot'	o-gúru	ɔ-gúru
		/-nə́mə/	'tongue'	o-nɨ́mɨ	ɔ-nɨ́mɨ

c.	/-sɛmɛ/	'grain'	ɔ-simi	a-simi
	/-gɔ̀nɔ/	'disease'	ɔ-gùnu	a-gùnu
	/-kara/	'rope'	ɔ-kiɾɨ	a-kiɾɨ

When occurring with the class 3 singular prefix /u-/, there are no complications: the vowel height of the root simply transfers to the prefix (and all root vowels are pronounced [+high]). However, when the prefix is plural class 6 /a-/, we observe that it is one step lower than the corresponding singular. This is obtained by fusing the transferred root vowel height with the lower vowel height of /a-/. (Although this would predict that /a-/ should be realized *ə-* or *ʌ-*, secondary processes modify these impermissible outputs to *o-*, *ɛ-*, and *ɔ-*.) The fact that there is a step-wise lowering of the prefix might suggest that /a-/ functions like a downstep marker operating on vowel height (Hyman 1988: 263). However, neither the Esimbi facts nor any other such scalar segmental process shows the properties of canonical tonal downstep: Thus, there is no language where hypothetical $^{\downarrow}i$, $^{\downarrow}u$ (perhaps pronounced [ɪ, ʊ] or [e, o]) contrast only after /i, u/, as $^{\downarrow}$H contrasts only after another H. There also is no "ceiling effect" on subsequent vowels, such that C$^{\downarrow}$iCe and C$^{\downarrow}$eCɛ are pronounced [CeCɛ] and [CɛCæ], respectively – and there certainly is no such effect on subsequent words in the phrase, as in the case of tonal downstep. Since tone and vowel height otherwise share properties, for example, their gradience along a single F0 vs. F1 dimension, the only conclusion to draw is that register effects such as downstep make tone qualitatively different from segmental features.

In fact, perhaps the most significant difference between tone and segmental phonology concerns the ability of tonal processes other than register to apply at long distances at the phrase level. Recall from (33) that Ndebele spreads, while Zulu shifts an H tone to word-antepenultimate position. In other Bantu languages an H tone spreads or shifts to a designated syllable in a following WORD. The words in the Shambala utterance in (50a) are all underlyingly toneless, and are therefore pronounced all L (Philippson 1998: 320):

(50)	a.	mawe	magana	mane	na= miloŋgo	mine	'440 stones'
	b.	magí̲	mágána	mat̲átú	ná= mílóŋgo	mine	'340 eggs'

However, in (50b) the two underlying H tones (originating on the underlined vowels of /magí/ 'eggs' and /matátu/ 'three') spread to the penultimate syllable of the following phonological word (or clitic group). Corresponding to the word-level spreading vs. shifting difference of Ndebele vs. Zulu in (33), the following Giryama examples show a long-distance rightward shift or displacement to the penultimate syllable of the following word (Philippson 1998: 321; cf. Kisseberth and Volk 2007):

(51)	a.	ku-tsol-a	ki-revu	'to choose a beard'	/-tsol-/	'choose'
	b.	ku-o̲n-a	ki-révu	'to see a beard'	/-ón-/	'see'

As in Shambala, the words in (51a) are both toneless, and the phrase is pronounced all L. In (51b), the verb root /-ón-/ 'see' carries an H tone which is displaced onto the penultimate of the toneless word /ki-revu/ 'beard'. It is examples such as these which motivate Yip (2002: 133) to remark that "the most striking property of African tone is its mobility". One looks in vain for a phrasal spreading or displacement of a segmental property in this way: Vowel-, consonant, and nasal harmonies, which appear to most closely mimic the effects of tone spreading, are typically limited to a word-size domain (which may include clitics). In the rare cases where vowel harmony hops over a word boundary, the affected target is typically a grammatical morpheme, as in Kinande, or the process instead represents a local coarticulatory effect, as in Nez Perce (Aoki 1966). In (51b) we have a robust case of an H tone shifting from one lexical word to another – something which is exactly duplicated in a number of other Bantu languages, for example, Digo (Kisseberth 1984: 163–164) and Zigula (Kenstowicz and Kisseberth 1990: 175).

While anticipatory spreading and shifting were said to be rare at the word level, phrasal anticipation appears to be quite natural. Consider the following forms from Tiriki [Kenya] (Paster and Kim 2007) in (52).

(52) /xu-molom-el-a/ 'to speak for' /xu-rhúmul-il-a/ 'to hit for'

a.	xù-mòlòm-èl-à mù-lìmì	xú-rhúmùl-ìl-à mù-lìmì	/mu-limi/ 'farmer'
b.	xú-mólóm-él-á mú-línà	xú-rhú↓múl-íl-á mú-línà	/mu-lína/ 'friend'

Both /xu-molom-el-a/ and /mu-limi/ are underlyingly toneless and pronounced all L in isolation, while /xu-rhúmul-il-a/ and /mu-lína/ have an underlying H on their first root syllable, which spreads onto the prefix: [xú-rhúmùl-ìl-à], [mú-línà]. In (52a), where the infinitive is followed by toneless /mu-limi/, no further change is observed. In (52b), however, the H of /mu-lína/ is anticipated not only onto its prefix /mu-/, but also onto all of the toneless TBUs of the preceding word. As seen, the toneless verb /xu-molom-el-a/ becomes all H, while the H of /mu-lína/ is anticipated up to the H of /xu-rhúmul-il-a/. Where the two H tones meet, we observe the indicated downstep. The same anticipatory process will apply through multiple words and phrases: /xu-rhúmulil-a + mu-limi + mu-lína/ → [xú-rhú↓múl-íl-á mú-límí mú-línà] 'to hit the friend for the farmer'.

Recall the Luganda H tone plateauing process which was illustrated word-internally in (40b). As seen in (53), the process actually applies across words within certain postlexical tonal domains (Hyman *et al.* 1987: 89):

(53) a. y-a-láb-à 'he saw' bi-kópò 'cups' by-aa= Walúsìimbi 'of Walusimbi'
H L H L H L

b. y-a-láb-á bí-kópó by-áá= Wálúsìimbi 'he saw the cups of Walusimbi'
H ------- H ---------- H L
L L
↓ ↓
Ø Ø

As indicated, each of the three words in (53a) has an H to L pitch drop. However, when combined in (53b), the whole sequence from the first to last H is pronounced on an H tone plateau. This is obtained by deleting Ls which occur between Hs, followed by plateauing.

As seen in the inputs in (53a), Luganda contrasts H, L, and Ø at an intermediate stage of representation (at the output of the word phonology). Any word-level Ls which occur between Hs are deleted within the proper tonal domain, no matter how distant the TBUs of the surrounding Hs may be. Another deletion process which takes place at a distance occurs in Peñoles Mixtec (Daly and Hyman 2007). In this language the underlying tones are /H/, /L/, and /Ø/, with an OCP(L) constraint triggering the following L tone deletion rule in (54a):

(54) a. L → Ø / L ___

b. ɨ̀ɨ^{N} dɨ̀i-ni-kwe-ši kada-kwe-ši ɨ̀ɨ^{N} ɨ́ɨ^{N} čɨ̀u^{N} → ɨ̀ɨ^{N} dɨ̀i-ni-kwe-ši kada-kwe-ši ɨ̀ɨ^{N} ɨ́ɨ^{N} čɨ̀u^{N}
one alone-only-pl-she ʀᴏᴛ.do-pl-she one one work L Ø
'Only one of them will do each of the jobs.'

This rule of L tone deletion applies across any number of toneless TBUs which may intervene. The rule thus applies to the second underlying /L/ in (53b), where there are 12 intervening toneless TBUs occurring between it and the preceding /L/. Like the cases of tone spreading, shifting, and plateauing, such extreme deletion at a distance is without parallel in segmental phonology.

While the above examples have to do with natural tonal processes which apply at a distance, another widespread phrasal phenomenon concerns cases where the tones of certain constructions are uniquely determined by the tones of the first morpheme or word. A well-known example concerns Shanghai compounds and other tightly bound constructions (Zee and Maddieson 1979: 109). As seen in the two examples in (55), this also is obtained by a two-step deletion + spreading process. All but the first tone is deleted, after which the second of the two features of the tonal contours is assigned to the second syllable. In the examples, the third syllable acquires a default L tone.

(55) a. 'illuminate' 'symbol' 'machine' 'camera'

tsɔ	+	ɕiã	+	tɕi	→	tsɔ	ɕiã	tɕi	→	tsɔ	ɕiã	tɕi
M H		M H		M H		M H				M	H	(L)

b. 'sky' 'studies' 'terrace' 'observatory'

t^{h}i	+	vəŋ	+	dɛ	→	t^{h}i	vəŋ	dɛ	→	t^{h}i	vəŋ	dɛ
H L		L H		L H		H L				H	L	(L)

While the tone melodies of languages such as Kpelle, seen earlier in (18), are normally a property of words, the Shanghai data show that tones may also be mapped over phrases. In fact, it is quite common for the tones of syntactically

conditioned constituents to be determined by the underlying tone or tone pattern of the first word. Based on their tonal behavior, Efere (2001b: 158–159) sets up the four following classes A–D in Izon (Bumo, Nigeria):

(56)

class	schema	tone pattern determined by the A–D class of the phrase-initial word
A	(L) H + H	all TBUs in the phrase = H
D	(L) H + HL	first word = all H, H spreads one TBU to the right; other TBUs = L
B	(L) H + L	first word = all H; subsequent TBUs = L
C	(L) HL + L	first word keeps its HL drop, remaining TBUs = L

The (L) in parentheses refers to the fact that vowel-initial words can begin H or L, whereas consonant-initial words begin H. Since only the tones of the first word are relevant, we can assume, as in Shanghai, that the tones of non-initial words are first deleted. The four classes are illustrated in (57) in the frame . . . /náná kímí/ 'man who owns/has . . .' (whose underlying H tones are deleted):

(57)

A	(L) H + H	bélɛ́	'pot(s)'	→	bélɛ́ náná kímí	(H spreading to end)
D	(L) H + HL	ìkíɛ́	'friend'	→	ìkíɛ́ nánà kìmì	(H spreading one TBU)
B	(L) H + L	wárí	'house'	→	wárí nànà kìmì	(no H spreading)
C	(L) HL + L	sɛ́rì	'scarf'	→	sɛ́rì nànà kìmì	(no H spreading)

As indicated, classes A, B, and D have a first word which is all H (class C words, which have a pitch drop, are largely borrowings). They differ by the degree to which this H affects subsequent words within the tonal phrase: all the way to the end (A), one TBU onto the second word (D), no spreading (B).

While Williamson (1988) uses a system of tonal melodies + floating tones to account for the variations in such tone patterns found in related Ijoid languages and dialects, at some point it becomes quite difficult to predict the phrasal patterns from specific underlying tones on the first word which determines them. As a case in point, in Urarina (Peru), tone is also determined by the first word of the phrase, which (Olawsky 2006) groups into the four classes A–D:

(58)

class	tone pattern determined by A–D class of the phrase-initial word
A	first word = L; H is assigned to initial syllable of following word
B	first word = L; H is assigned to 3rd syllable of following word
C	first word = L; H is assigned to last syllable of final word of phrase
D	first word keeps its H tone when a word follows, all the rest = L

In isolation, Urarina words generally have a single final H. Whereas the first word is all H in three out of the four classes in Izon, in Urarina the first word of a phrase is L except in class D (which also includes some words that have their H on the penultimate rather than final syllable). As summarized above and illustrated

when preceding the trisyllabic verb *rʉ.a.kaa* 'carries/carried (3sg.)' below, the difference concerns the placement of an H tone on the next word:

(59)	A raaná	'peccary (sp.)'	→ raana rʉ́.a.kaa	'he has carried peccary'
	B obaná	'peccary (sp.)'	→ obana rʉ.a.káa	'he has carried peccary'
	C reemaé	'dog'	→ reemae rʉ.a.kaá	'he has carried dog'
	D makusajarí	'pepper'	→ makusajarí rʉ.a.kaa	'he has carried the pepper'

The above represents the general case. One complication is that class B will assign the H to the second syllable of a trisyllabic word whose last syllable is short. It is possible to treat class C as assigning a toneless pattern (with the phrase-final mora being due to a default rule) and class D as borrowings and exceptions. This would still leave the difference between class A and class B. One is tempted to start with an H assigned to one syllable which then shifts two or three syllables to the right, but there does not appear to be any evidence for this, hence Olawsky's class A–D approach.

The same conclusion is suggested by the facts from the Move dialect of Yagaria (Papua New Guinea). As indicated in (60), according to Ford (1993: 196–197) words either have stable (S) tones or belong to one of three "unstable" tone classes (U_1–U_3):

(60)	a.	S	hógà	'left'	→ hógà kàyàlè	'left pig'	(no change + all L)
		S	fáipái	'white'	→ fáipái kàyàlè	'white pig'	(no change + all L)
	b.	U_1	lòlé	'two'	→ lòlè kàyàlé	'two pigs'	(all L + L-L-H)
		U_3	fèlá	'wild'	→ fèlà kàyálé	'wild pig'	(all L + L-H-H)
		U_2	kòlí	'scared'	→ kòlì káyálé	'scared pig'	(all L + H-H-H)

In the above examples the second word is /kàyálè/ 'pig', which becomes all L after a stable tone word, as in (60a). In (60b), all three unstable tone words have L-H tone in isolation. As seen, they have different effects on the next word: U_1 places a single H on the last syllable, U_2 assigns H to the last two syllables, and U_3 assigns H to all three syllables. One interpretation might be to view the second word as becoming toneless after a stable tone word (and hence all L). Unstable tone words would lose their H tone, and assign a single H to one syllable of the next word, which then spreads onto following syllables. In this case, we would obtain /kayalé/ (after U_1), /kayále/ (after U_2), and /káyale/ (after U_3). It is not clear how these different H tone placements could be predicted from different underlying representations on the first word. While the assignment of a single H to the next word in Urarina and Yagaria might suggest "accent-like" behavior, it is important to note that stress-accent systems do not appear to show such effects. Thus, there appears to be no language where class A words assign a final stress to the next word, class B a penultimate stress, and class C an initial stress. Rather, phrase-level arbitrary classes seem to be a tonal phenomenon.

The final evidence that more is going on than a simple mapping from the first word to the phrase is found in Wuxi, a Northern Wu dialect that has been studied

in detail by Chan and Ren (1989). The following table is reproduced from Chen (2000: 323) (cf. also Yip 1989: 156–159):

(61)

1st σ			2nd σ		
			T3, 4, 5, 6	T1, 2	T7, 8
A LLH	T3	→	C LHH		
	T6	→	D LHL		
B HHL	T1	→	A LLH		
	T4	→	A LLH	D LHL	
	T7q	→	A LLH	D LHL	B HHL
C LHH	T5	→	B HHL		
D LHL	T2	→	B HHL		
	T8q	→	A LLH	D LHL	

Synchronic Wuxi contrasts four surface tones (A–D), which, as indicated, are characterized in terms of four tone features (which Yip 1989 reanalyzes, respectively, as L+LH, LH+L, H and L+HL, where LH and HL are tonal contours in the sense of (21a)). These tones derive from various mergers of the Middle Chinese eight categories (T1–T8), where T7q and T8q refer to the two tones on stopped syllables. Wuxi, like Shanghai, maps a single tone pattern over a phrasal domain. However, as seen in (61), the exact shape of the pattern is determined not only by A–D identity of the first tone, which represents mergers of the Middle Chinese tones, but also by a three-way contrast in the historical identity of the second tone. There thus is considerable arbitrariness from a synchronic point of view. As Chan and Ren (1989) point out, Wuxi tone sandhi was originally right-dominant (whereby the first tone changes), but then became left-dominant, deleting the second tone, as in Shanghai. Chen (2000: 325) provides the following sample derivation to illustrate:

(62) a. na + dã ‘milk candy’
HHL LHL (base tones)
LHL LHL (first tone undergoes sandhi)
b. LHL Ø (deletion of second tone)
LH HL (tone spread)

In (62a) the initial HHL (B) tone changes into an LHL (D) tone. This is followed in (62b) by first deleting the LHL (D) tone of the second morpheme, which then

allows the one LHL tone to map over the two syllables. The result is an alternation between the /HHL/ of the first syllable with a /LHL/ which it maps over the two syllables. The Wuxi situation demonstrates the extent to which synchronic tonal properties can encode the history of a language, particularly at the phrase level. It would be extremely interesting to know the historical origins of the Izon, Urarina, and Yagaria systems.

We have been concerned with the fact that while a phrase-initial word can assign an arbitrary tone to a following word, this does not seem to be true either of segmental phonology or of stress. There is no language where, say, the first word arbitrarily assigns a [+nasal] to different syllables of the next word, nor are there stress classes that assign different stresses to a following word. It should be noted that the examples cited above are all head-final, such that it is an initial modifier or complement which determines the tonal properties of the phrase. This contrasts with grammatical feature assignments which normally go from head to dependent. In this context consider the highly unusual tonal agreement from Barasana (Colombia) (Gomez-Imbert and Kenstowicz 2000: 438–439), which has an underlying contrast between bimoraic morphemes which are all H vs. HL. As seen in (63), possessive pronouns condition tonal agreement on the following noun:

(63)

	~kúbú (H) 'shaman'	~bídì (HL) 'pet'
~bádí (H) 'our'	~bádí ~kúbú	~bádí ~bídí
~ídà (HL) 'their'	~ídà ~kúbù	~ídà ~bídì

When the pronoun is H, as in the case of *~bádí* 'our', *~bídì* 'pet' changes to H. Similarly, when the possessive pronoun is HL, as in the case of *~ídà* 'their', *~kúbú* 'shaman' changes to HL. Since agreement is normally a grammatical concept, one might be tempted to consider the above facts not as tonal agreement, but as grammatical agreement (where H and HL are exponents of grammatical features). However, the agreement in Barasana goes from modifier to head, hence again, in the wrong direction. It is tempting to instead see the examples in (55)–(63) as cases where the head (and potentially other non-initial words) undergoes reduction, followed by additional processes of tone spreading, tone assignment, tone agreement and so on.

To summarize this section, we have seen that tone is capable of reaching deeply across word boundaries for both grounded and not-so-grounded processes, for example, downstep anticipation, non-local H spreading, shifting and plateauing, OCP effects, tone mapping. Segmental phonology does not have such a long reach, but rather is restricted to local adjustments, as when the last segment of one word interacts with the first segment of the next. Stress offers more possibilities, particularly when the stress of one word is retracted or deleted so as to avoid a clash with the stress of another word. It too, however, does not show all of the parallels

illustrated in the above discussion. If tone is really different, then these findings have the following implication for word-prosodic typology: any system that does what only tone can do is tone. The issue of typology is taken up in the following section.

5 Tone vs. Accent

While I have treated each of the phenomena and all of the examples in the preceding sections as strictly tonal, there has been a tendency to view languages which restrict the distribution of their tones as "accentual." For example, since Urarina assigns a single H within the appropriate phrasal domain, might this H be a "pitch accent?" However, since most of the properties discussed in Section 4.2 do not have analogues in stress-accent systems, their identification as accentual phenomena is not obvious. In this section, I will argue that the phenomena in question are typical of tone systems, particularly those which place restrictions on the distribution of their tones.

Within the generative tradition, the study of word-prosodic typology was greatly influenced by McCawley (1968b, 1970), who attempted to set up a principled distinction between tone vs. pitch-accent systems based both on distributional properties and rule types (tones tend to assimilate; accents tend to dissimilate or reduce). A survey of subsequent literature reveals that the terms "accent," "pitch accent," and "tonal accent" have generally been used to refer to tone systems which are defective in the sense of restricting tones by number of contrasts or by position: "A pitch-accent system is one in which pitch is the primary correlate of prominence and there are significant constraints on the pitch patterns for words . . ." (Bybee *et al.* 1998: 277). Among such "significant constraints" are those enumerated in (64), where the tone in question is most commonly /H/:

(64) A tone may be . . .

a. obligatory: "at least one" must occur per domain (e.g. word)
b. culminative: "at most one" can occur per domain
c. privative: the underlying contrast is between presence vs. absence of the tone
d. predictable: assigned to positions by rule
e. restricted: occurring in only in certain positions (stressed syllable, first two syllables)
f. reducible: subject to reduction, subordination (e.g. in compounding, defocusing)

However, most or all of the above properties can be found in unambiguous tone systems. For example, consider OBLIGATORINESS as reflected in the attested tone patterns of Chuave (Papua New Guinea) (Swick 1966; Donohue 1997: 355) in (65).

(65)

scheme:	/H/	/HL/	/LH/	/HLH/	/LHL/
1σ = 1	H				
2σ = 3	H-H	H-L	L-H		
3σ = 7	H-H-H	H-H-L H-L-L	L-H-H L-L-H	H-L-H	L-H-L
4σ = 15	H-H-H-H	H-H-H-L H-H-L-L H-L-L-L	L-H-H-H L-L-H-H L-L-L-H	H-H-L-H H-L-H-H H-L-L-H	L-H-H-L L-H-L-L L-L-H-L

As seen, all combinations of H and L tone are found on words of one to four syllables, except an all L pattern. H tone is therefore "obligatory," but hardly accentual, given, for example, the 15 patterns possible on four-syllable words. Numerous other tone systems have two, three, or four word-tone patterns requiring an H, for example, /H, LH/ in Hup (Colombia) (Epps 2005), /H, HL, LH/ in Dom (Papua New Guinea) (Chida 2001), /H, HL, LH, LHL/ in Dogon (Jamsay) (Mali) (Heath 2008). It is not clear that there is anything special about such obligatory-H systems vs. those which allow a /L/ pattern. In fact, the distributions in (66) suggest that the obligatoriness of H in Tanimuca (Colombia) (Keller 1999) may be accidental:

(66)

2σ = 3	H-H	H-L	L-H				*L-L
3σ = 6	H-H-H	H-H-L	H-L-H	L-L-H	L-H-H	L-H-L	*L-L-L, *H-L-L

All patterns occur on bisyllabic words except L-L, suggesting obligatory H. However, three-syllable words show only six patterns. Obligatoriness predicts that L-L-L should not be possible, but has nothing to say about the absence of H-L-L. The generalization, therefore, may be that a word cannot end in two L tones. Unfortunately there are no monosyllabic words in the language, which could disambiguate between the two interpretations. In any case, the patterns which do contrast on three syllables again argue for a tonal rather than accentual interpretation.

While obligatory H is much more common, there are also cases of obligatory L: in Hakha Lai (Burma), the three underlying tones are /HL/, /LH/, and /L/ (Section 2). Not only does /H/ not exist, but all of the tonal alternations conspire to preserve input L features (Hyman and VanBik 2004). A number of languages have a tonal contrast on only one (possibly stressed) syllable per word. Thus, Dadibi (Macdonald and Macdonald 1974) and Fasu (May and Loeweke 1964), both spoken in the Southern Highlands Province of Papua New Guinea, require one syllable per word to be /H/ or /L/, remaining syllables being toneless, while Pame (Central, (Mexico) (Gibson 1956), has one obligatory /H/, /HL/, or /L/ syllable per word.

It appears that tone is the only phonological feature which can have this obligatory property. (In the case of stress, it is the metrical structure that is obligatory.) Although all phonological systems have at least two contrastive vowel heights (Maddieson 1997: 636), no language requires every word to have at least one high vowel. Similarly, all languages have oral stops, but unless Rotokas is analyzed as /p, t, k, b, d, g/ rather than Firchow and Firchow's (1969) /p, t, k, β, ɾ, g/, no language requires every word to contain at least one stop. Although tone is once again different from segmental phonology, obligatory H or L does not mean that a tone is an accent.

The next restrictive property on tone is CULMINATIVITY: in some languages tone is not obligatory, but is restricted to at most one occurrence per domain. The standard example here is Tokyo Japanese, which, as seen in (67), has been subject to both accentual and tonal analyses (McCawley 1978; Haraguchi 1979; Poser 1984, Pierrehumbert and Beckman 1988, etc.):

(67)

		'pillow' + nom.	'heart' + nom.	'head' + nom.	'fish' + nom.
a.	accentual	ma↓kura ga	koko↓ro ga	atama↓ga	sakana ga
b.	tonal	makura ga | H L	kokoro ga | H L	atama ga | H L	sakana ga
c.	approx. pitch	mákùrà gà	kòkórò gà	àtāmá gà	sàkānā gā

Tokyo Japanese has been characterized as having at most one abstract accent (*) per word, one HL pitch drop, or one prelinked H tone. Culminative /H/ occurs in Somali (Hyman 1981; Saeed 1999) and Ocaina (Peru) (Agnew and Pike 1957), and is particularly well-attested in Papua New Guinea, for example, in Bahinemo (Dye 1992), Baruya (Lloyd and Lloyd 1992), Tinputz (Hostetler and Hostetler 1975), and Wantoat (Davis 1969). As in the case of obligatoriness, there are variations on the identity of the culminative tone(s): /HL/ (vs. /H, L/) in Obukuitai (Indonesia, Papua) (Jenison and Jenison 1991), /H/ or /HL/ (vs. /Ø/) in Awad Bing (Papua New Guinea) (Cahill 2001), /H/, /HL/, or /LH/ (vs. /L/) in Puinave (Colombia) (Girón Higuita and Wetzels 2007).

While it was pointed out that only stress and tone can be obligatory, a number of other features have been reported to be culminative, for example, aspiration and glottalization in Cuzco Quechua (Parker 1997: 2), vowel length in Mam (Willard (2004: 7), mid vowels in Punu (Kwenzi Mikala 1980: 8; Fontaney 1980b: 55), nasalized vowels in Karo (Gabas 1999: 42n). Since one does not speak of "mid vowel accent" or "nasal accent," and so on, we can assume that culminativity is simply a restriction that can be placed on tone, as it can be on other features. That this is the right move is seen from Donohue's (1997: 367–368) presentation of the tone patterns in Arigibi Kiwai (Papua New Guinea) in (68).

(68)

schema	1σ	2σ	3σ	4σ
/L/	L	L-L	L-L-L	L-L-L-L
/HL/		H-L	H-L-L	H-L-L-L
/LH/		L-H	L-L-H	L-L-L-H
/LHL/			L-H-L	L-H-L-L L-L-H-L

As seen, /L/ is obligatory, but /H/ is culminative. If these are taken to be diagnostics for accent, then the question would be: Which is the accent? On the other hand, /H/ is both obligatory and culminative in languages such as Kinga (Schadeberg 1973), Una (Indonesia, Papua) (Donohue 1997), and Nubi (Uganda) (Gussenhoven 2006). In such cases, since only pitch is involved in marking the obligatory+culminative prominence, it is not clear whether the H tone should be viewed in metrical vs. strictly tonal terms. In Kinga, at least, where the /H/ is assigned by mora, thereby creating syllables with a [HL] vs. [LH] contrast, as in Somali in (3b), this would be an important difference with canonical metrical systems where the syllable is the stress-bearing unit.

The remaining properties in (64) are probably even less reliable indicators of an accentual system, although they have historically played into such analyses. Some of the studies in Clements and Goldsmith (1984), van der Hulst and Smith (1988), and Inkelas and Zec (1990) deal with Bantu systems which have a privative /H/ vs. /Ø/ opposition (cf. Odden 1988, 1999). One such system is Haya, where one might also be misled by the fact that nouns in isolation can bear at most one H or HL tone, for example *o-mu-zi* 'root', *o-mú-ti* 'tree', *o-mu-limi* 'farmer', *o-bugólo* 'snuff', *o-mu-kâzi* 'woman'. However, as seen in (69), the word /ó-mu-tí/ 'tree' has two underlying H tones and can be realized with any of the eight possible combinations of H and L (< /Ø/) tones (Hyman and Byarushengo 1984: 56):

(69)

a.	L-L-L	:	o-mu-ti gwaa káto	'Kato's tree'
	L-L-H	:	o-mu-tí gwange	'my tree'
	L-H-L	:	o-mú-ti	'tree'
	L-H-H	:	o-mú-tí káto	'a tree, Kato!'
b.	H-L-L	:	a-gul' ó-mu-ti gwaa káto	'he buys Kato's tree'
	H-L-H	:	a-gul' ó-mu-tí gwange	'he buys my tree'
	H-H-L	:	a-gul' ó-mú-ti	'he buys a tree'
	H-H-H	:	a-gul' ó-mú-tí káto	'he buys a tree, Kato!'

The rules involved are as follows: (i) An H tone vowel becomes L after pause, as in (69a); (ii) an H tone is deleted before a genitive noun phrase; (iii) a phrase-final H tone is anticipated onto the penultimate (e.g. before vocative 'Kato!'); (iii) an H tone becomes L before pause (where it also will be phrase-final). (The apostrophe in the examples in (69b) indicates the elision of the final vowel of a-gul-a 'he buys'.) Even though the underlying /H/ vs. /Ø/ contrast is privative, the system is clearly

tonal. The same is true of systems where Hs are attracted to metrically strong positions, for example, the penultimate in Shambala (50b) and Giryama (51b).

Finally, bona fide tone systems may also show phrasal reduction phenomena reminiscent of loss of stress when out of focus or stress subordination in compounding. For example, the East Sudanic language Nara (Eritrea) distinguishes at least the tone patterns in (70), taken from Hayward (2000: 255), where the mora (μ) is a vowel or sonorant:

(70)

schema	μ	μμ	μμμ	μ-μ	μμ-μ	μ-μμ	μ-μ-μ	μμ-μ-μ(μ)
/L/	(L)	LL	(LLL)	(L-L)	LL-L	(L-LL)	L-L-L	LL-L-L
/H/	H	HH	(HHH)	H-H	HH-H	(H-HH)	H-H-H	(HH-H-H)
/LH/		LH	LLH	L-H	LL-H	L-LH, L-HH		LL-L-H
/HL/	HL	HL	HLL	H-HL	HL-L, HH-L	H-LL	H-H-L	
/LHL/			LHL		LH-L	L-HL	L-H-L	LL-L-HL

As seen, the five schemas establish that there is a full tone system, although with the following constraint: ". . . attention is directed towards an obvious asymmetry with regard to the distribution of the two tones, for we do not find an HLH contour as a counterpart to the LHL contour . . ." (Hayward 2000: 256). Recall that /HLH/ was also missing from Kpelle in (18), and that this sequence served as the input to H tone plateauing in Luganda in (40b) and (53b). In fact, as summarized in the table in (71), whenever two words occur which produce an H-L-H sequence in a COMPLEMENT+HEAD construction, the second H becomes L, as indicated by L in the appropriate cells:

(71)

$W_1\downarrow/W_2 \rightarrow$	/L/	/H/	/LH/	/HL/	/LHL/
/L/					
/H/			L		L
/LH/			L		L
/HL/		L	L	L	L
/LHL/		L	L	L	L

Thus, H-H # L-H will become H-H # L-L, H-L # H-H will become H-L # L-L, and so forth. Whereas Luganda changes H-L-H to H-H-H, Nara modifies H-L-H to H-L-L. Both are a response to the constraint against *HLH (Cahill 2007) or *TROUGH (Yip 2002: 137). Since the relevant Luganda constructions have the reverse

HEAD+COMPLEMENT structure, the two languages reveal that it is the tones of the head that are modified to avoid a tonal trough. While this corresponds with the observation that the focus position is often post-verbal in a VO language, but pre-verbal in an OV language (Harries-Delisle 1978: 464), we do not need to see the phenomenon as accentual. Rather, prosodic features such as stress, tone, and also vowel length (Kisseberth and Abasheikh 1974) can be sensitive to specific syntactic configurations.

To summarize this section, we have seen that tones can be relatively free or considerably restricted. At one end of the spectrum, tone specifications are "dense," with a tone feature required for every TBU. At the other end, tone specification is "sparse," with tone features subject to obligatoriness, culminativity, privativity, or restriction to specific positions. While systems which tend toward the latter characterization have sometimes been termed "accentual," this is an analytical move which is hard to justify on independent grounds. As Gussenhoven (2004: 42) puts it, "'Accent' . . . is an analytical notion and cannot be measured. [It is] thus different from stress, which is typically an observable phenomenon, and different also from tone, whose existence is equally measurable." Thus, while some languages must be analyzed with stress and others with tone, it is important to underscore that no language MUST be analyzed as "pitch accent." A tonal interpretation is always possible.

The greatest challenge to this position comes from tone systems which have a metrical character. Consider, for example, the case of the Iroquoian language Seneca (Iroquoian) (Chafe 1977, 1996). Seneca and closely related Cayuga have received considerable attention in the metrical literature (see Hayes 1995a: 222–226 and references cited therein), which has generally analyzed Seneca as having iterative iambs constructed left-to-right. The properties of Seneca in (72) summarize the trochaic reanalysis in Melinger (2002):

(72) a. mark the first syllable extrametrical
b. build bisyllabic trochees left-to-right
c. assign an H tone to the first syllable of a trochee iff either syllable is closed

As indicated, Seneca has a metrical system. However, several things set it apart from more usual stress systems. First, the trochees are not used to establish stress, but to assign H tones. Second, the H tones are assigned by open vs. closed syllable. Third, the required closed syllable can be in either position of the trochee. This produces the rather unusual tonal distributions seen in the schemas in (73).

(73) a. <σ> (CáC.Ca) (CaCa)
b. <σ> (Cá.CaC) (CaCa)
c. <σ> (Cá.CaC) (CáC.Ca)
d. <σ> (Ca.Ca) (Ca.Ca)

In each example, the first syllable has been marked off as extrametrical, and the following four syllables have been grouped into two trochees. In (73a) the head

of the first trochee consists of a closed syllable and thus receives an H tone (vs. the second trochee which consists of two open syllables). (73b) also receives an H on the head of the first trochee, although in this case it is the closed syllable of the non-head which licenses the H. In (73c) both trochees contain a closed syllable, and so an H tone is assigned to the first syllable of both, in violation of culminativity. Finally, since there are no closed syllables in (73d), no H tone is assigned, in violation of obligatoriness. The question here is whether Seneca has a stress system, a tone system, both, or something else. Following Prince (1983), Hayes suggests that "Lake Iroquoian accent is partly metrical, partly tonal" (Hayes (1995a: 225). In other words, Seneca has a stress system, which happens to be realized in terms of non-obligatory, non-culminative H tone (as well as by other means, for example, lengthening of a strong open-syllable penultimate vowel). Seneca thus offers a case of where tone is superimposed on stress (Hyman 1978: 5), something which might be designated as "a metrical tone system" for convenience. With time, the subsequent segmental changes, which have already begun to obscure the transparency of the metrical system, may make the placement of H tones less predictable. In any case, we are far from the phenomena which are called "accent" because of the sparseness of their H tones. The general conclusion of this section is that so-called pitch accent has no independent properties, rather represents a "pick-and-choose" between a number of properties which are normally associated with stress vs. tone systems (Hyman 2001b, c; 2009a). The languages cited in this section are thus better viewed as "restricted tone systems" (Voorhoeve 1973) rather than belonging to a coherent third category.

6 Conclusion

In the preceding sections I have covered some of the conceptual and analytical problems one faces in approaching the study of tone. As we have seen, tone systems are both complex and diverse. The impression one gets is that Hs, Ls, Ms etc., may exhibit as much variation as all of segmental phonology combined: *Tone can do everything that segmental or metrical phonology can do, but the reverse is not true*. Although we rejected the misconceptions about tone in Section 2, the conclusion nonetheless is that tone is both quantitatively and qualitatively different from segmental features and from the other two traditional suprasegmentals (length and stress). Compared with other phonological properties, we can say the following about tone:

(i) *Tone is the most syntagmatic:* as we saw especially in the phrasal examples in Section 4.2, tone shows the greatest tendency either to wander from its original TBU (Yip's "mobility") or to have effects with other tones at great distance. Such properties are particularly observed in systems which contrast tonally marked vs. /Ø/ TBUs. As few if any other features show such postlexical action at a distance, it is not clear what tone's closest competitor is.

(ii) *Tone is the most paradigmatic:* despite the highly syntagmatic nature of tone, tone can also be highly paradigmatic, especially in languages in which monosyllabic languages contrast multiple tone levels and tonal contours on monosyllabic words (recall Trique (Itunyoso) in (9). No other single phonetic dimension offers as many potential phonological contrasts along a single dimension as F0 does for tone – up to five tone heights, as was exemplified from Kam (Shidong) in (7). The only close competitor is vowel height. Among the 451 languages in the UPSID database (Maddieson and Precoda 1990) the only language which has the five-way contrast /i, ɪ, e, ɛ, æ/ in its inventory of short front unrounded oral vowels is Somali. However, it is clear that this is made possible by the fact that Somali has two sets of vowels which harmonize: peripheral /i, e, æ, u, o/ vs. centralized /ɪ, ɛ, a, ʉ, ö/ (Saeed 1993). This suggests that another phonetic dimension, perhaps [±ATR], is involved. While some tone systems show that the multiple levels naturally group in higher vs. lower registers (Yip 1980), sometimes with accompanying phonation differences (breathiness, creakiness), in many multilevel tone systems, the only differences are in F0. The possibility for multiple paradigmatic contrasts on a single syllable, including tonal contours, is the aspect of tone that makes it the most different from the use of pitch to mark stress and intonation.

(iii) *Tone is the most analytically open-ended*: tone appears to offer a wider range of reasonable analyses and interpretations than other phonological features. In different languages, a two-level tone system can be analyzed as /H, L/, /H, Ø/, /Ø, L/, or /H, L, Ø/, and there are more options if one succumbs to the temptation to treat tone as "accent" (Section 5). Similarly, depending on the language, the M of a three-level system can be a lowered H, a raised L, an independent third tone, or /Ø/. Using the tone features introduced by Yip (1980), M can be [+upper, –raised] or [–upper, +raised]. While this feature difference is quite useful when there are four contrastive tone levels, even a three-level tone system can have two phonetically M tones which contrast in certain contexts, suggesting that they are underlyingly different. While it has been occasionally claimed (e.g. by Goldsmith 1985) that other features can be binary in some languages but function privatively in others, or that one feature value can be marked in one language vs. unmarked in another (Hume 2003, 2007), tone offers especially compelling examples of both. Thus, while /H/ vs. /Ø/ is more common than /L/ vs. /Ø/, both are found in Bantu (Hyman 2001b) and in Athabaskan (Hargus and Rice 2005), and "marked L" is also found in Malinke (Kita) ((Creissels and Grégoire 1993), Galo (NE India) (Post 2007), and the closely related NW Amazonian languages Bora (Weber and Thiesen 2000) and Miraña (Seifart 2005).

(iv) *Tone is the most abstract*: in many cases the relation between input and output tones is a very abstract or indirect one. In the case of two-level tone systems, this is in part due to the considerable possibilities for spreading, shifting, deletion, and insertion, as when the tones of Haya /ó-mu-tí/ 'tree' appeared to become inverted to [ò-mú-tì] in (69) above. In addition, underlying systems can be realized with more levels in the output than they started with. Thus, although Ngamambo only has underlying /H, L/, after an H, there is a five-way surface

contrast between H, M, ↓M, L and L°, the last being a level L tone distinct from an L which falls in pitch before pause (Hyman 1986). Floating tones add considerably to the possibilities for abstract underlying representations, and are often easier to motivate than their segmental counterparts.

(v) *Tone is the most arbitrary*: while much of tonology is phonetically grounded in a transparent way, many tonal alternations appear arbitrary from a synchronic point of view. This occurs both in morphology, for example, nominative vs. accusative tone marking in Maasai in (14), as well as at the phrase level, for example, the Wuxi case discussed in Section 4.2. In many cases one tone pattern simply corresponds arbitrarily to another. Perhaps the most famous case of this occurs involves Xiamen tone sandhi rule in (74a) (Chen 1987):

(74) a. T → T′ / ___ T within a tone group (T = base tone; T′ = sandhi tone)
b. 24, 44 → 22 → 21 → 53 → 44
c. # yi kiong-kiong kio gua ke k'uah puah tiam-tsing ku ts'eq #

44	24	24	21	53	44	21	21	53	44	53	32	= T
22	22	22	53	44	22	53	53	44	22	44		= T′

he by force cause I more read half hour long book
'he insisted that I read for another half an hour' (Chen 1987: 113)

While the chain shift in (74b) shows that each of the five non-stopped base tones is replaced by another tone, the actual featural changes cannot be motivated synchronically. It hard to think of an equally complex arbitrary input-output relation outside tone.

(vi) *Tone is the most autosegmental*: As was seen in Section 3, tone is THE autosegmental property par excellence. Compared to segmental features, tone is far more likely to float as a lexical or grammatical tone, to show stability effects, to undergo dislocation, or to interact with like features at a distance. Tone sequences are much more likely to be treated as contours which can be manipulated as units or as "melodies" which can be mapped over multiple TBUs. While there are analogues to many of the autosegmental properties of tone, tone provided – and still provides – the model.

To sum up the above observations in one sentence, tone is extraordinarily versatile, a lot of things at once. Over the past several decades there has been tremendous progress both in documenting this versatility and in developing models to deal with it. Autosegmental phonology has been successful largely, because, as Kenstowicz (1994: 312) puts it, ". . . tone behaves independently from other features and so can be discussed in relative isolation." However, it is when we approach the INTERDEPENDENCY of tone with other features that the picture becomes less clear. Although of great practicality, the use of informal Hs, Ls, Cs, and Vs disguises several interrelated questions which have not been settled: (i) What is or can be a TBU? The syllable? mora? segment within a rime? onset? (ii) Where does tone link up within the feature geometry? Exclusively to the TBU? In addition or instead of the laryngeal node, where Halle and Stevens' (1971 pitch-affecting glottal features [stiff], [slack], [constricted], and [spread] can naturally

interact with tone? (iii) What is the correct set of tonal features? Is a unique set of tonal features even possible, or does it vary from language to language? Possible solutions to such questions are considered in several monographs and dissertations, for example, Duanmu (1990), Bao (1999), Bradshaw (1999), Snider (1999), Yip (2002), and Pearce (2007b). The answers ultimately also bear on the question of how laryngeal gestures are implicated in the process of tonogenesis (Matisoff 1973; Hombert, Ohala, and Ewan 1979; Thurgood 2002). However, despite the large number of proposals, the field is far from consensus. Recently, Clements, Michaud, and Patin (2009), and Hyman (2009b) have argued that the kinds of arguments supporting a universal set of segmental distinctive features are largely absent or ambiguous when applied to tones. While the Chomsky and Halle (1968) feature specifications [+high, –low] were designed to capture both the intersecting class of high vowels, /i, ü, ɯ, u/ and their relation to palatal and velar consonants, the same cannot be said for [+upper, +raised] or [+stiff, –slack]. As opposed to [+high, –low], such features specify only one tone, namely H, not a class of tones, and also fail to capture five contrasting tone heights. Where they do interact with laryngeal features, for example, voicing, aspiration, or glottalization, the diachronic and synchronic correlations are often contradictory (see Kingston 2003, 2005 and references cited therein). As a result, tone is less reliably "gridded in" than other features, thus again suggesting that it is "different."

Given the independence and extreme range of tonal phenomena, perhaps there will not be a single, definitive answer to one or more of the questions raised in this chapter. In fact, the above questions (i)–(iii) may even contain a misunderstanding that we still have either about tone, or more likely, about phonology in general. Perhaps when these questions are answered, tone will once again contribute in a central way to phonological theory.

ACKNOWLEDGMENTS

I am grateful to various audiences who had heard and responded to different sections of this chapter under titles such as "Is tone different?" and "There is no pitch-accent prototype." I would especially like to thank Nick Clements, Grev Corbett, Denis Creissels, John Goldsmith, Carlos Gussenhoven, Morris Halle, Will Leben, Doug Pulleyblank, and Donca Steriade for their stimulating reactions, which still keep me thinking.

8 Harmony Systems

SHARON ROSE AND RACHEL WALKER

1 Introduction

This chapter addresses *harmony systems*, a term which encompasses consonant harmony, vowel harmony, and vowel-consonant harmony. Harmony refers to phonological assimilation for harmonic feature(s) that may operate over a string of multiple segments. This can be construed in one of two ways. Two segments may interact "at a distance" across at least one (apparently) unaffected segment, as shown for consonant harmony in (1a). Or, a continuous string of segments may be involved in the assimilation, as shown for vowel-consonant harmony in (1b). The subscripts refer to features or feature sets.

(1) a. *distance harmony*
consonant harmony $C_x\ V_y\ C_z \rightarrow C_z\ V_y\ C_z$
b. *continuous harmony*
vowel-consonant harmony $C_x\ V_y\ C_z \rightarrow C_z\ V_z\ C_z$

Although only three segments are represented in the diagram in (1), harmony can apply to longer strings. As for vowel harmony, it can operate at a distance depending on how one construes intervening consonants and/or vowels that are apparently unaffected by the assimilation. It may also be construed as continuous if intervening segments participate in harmony. Furthermore, vowel-consonant harmony can operate at a distance, skipping over some segments.

In this chapter, we first provide a descriptive overview of the basic patterns of the harmony systems outlined in (1), with a focus on the triggers (segments that cause harmony) and targets (segments that undergo harmony). We also touch on blocking segments (ones that halt harmony) and transparent segments (ones that

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

appear to be skipped by harmony but do not prevent it from extending past them). We then elucidate the main analytical trends and advances that phonological theory has brought to bear on harmonic systems. First, not all harmony systems show the same characteristics or are amenable to the same type of basic analysis. Specifically, there appears to be a split between consonant harmony on the one hand and vowel harmony and vowel-consonant harmony on the other. Second, there has been a shift from emphasis on tier-based representational solutions for issues such as blocking and harmony drivers in favor of alternative explanations articulated within Optimality Theory, such as phonetically-motivated featural co-occurrence constraints and agreement constraints that are non-specific about targets. Third, an increased appeal to functional explanations has been sought for harmony patterns. Finally, broader typological coverage has led to progress on topics such as directionality and consonant harmony, but has also provided challenges to well-established conceptual issues.

The chapter is organized as follows. In Section 2, we introduce and illustrate the main properties of harmony patterns. In Section 3.1 we discuss autosegmental approaches. In Section 3.2, we show how consonant harmony has come to be analyzed through correspondence-based constraints which require participating similar segments to match. Vowel harmony and vowel-consonant harmony share many issues, addressed in Section 3.3. These include conceptualization of the harmony imperative, or what drives harmony, feature classes, blocking and transparency. In Section 3.4 we discuss the disparate cases of non-local vowel-consonant harmony. In Section 4.1, we address directionality, and in Section 4.2, phonologically and morphologically defined harmony domains are explored. In Section 5 new trends and directions for future research are presented.

2 Harmony Patterns

2.1 Consonant Harmony

We define consonant harmony as assimilation between consonants for a particular articulatory or acoustic property operating at a distance over at least another segment.[1] Consonant harmony can involve both alternations in affixes and morpheme structure constraints (Shaw 1991; Hansson 2001b, 2010; Rose and Walker 2004).

2.1.1 Coronal Harmony The most commonly attested type of consonant harmony is *sibilant harmony*, which requires sibilant coronal fricatives and affricates to match for tongue tip/blade posture and location. It is widely attested in Native American languages, but also occurs elsewhere. In Ts'amakko, a Cushitic language of Ethiopia (Savà 2005), the causative suffix *-as* (2a) is realized as [aʃ] when palatoalveolar fricatives or affricates appear in the preceding stem (2b):

(2)	a.	ɓaɗ	'to hide'	ɓaɗ-as	'to make hide'
		ɠabb	'to take'	ɠabb-as	'to make take'

	bas	'to do'	bas-as	'to make sb. do'
	zaq'	'to slaughter'	zaq'-as	'to make sb. slaughter'
b.	ʧ'ur	'to throw'	ʧ'ur-aʃ	'to make sb. throw'
	ʃukuy	'to be scared'	ʃukuy-aʃ	'to scare'

Sibilant harmony operates across vowels and non-sibilant consonants, including other coronals. In (2b), the intervening segments do not block and do not participate in the harmony. In roots, in addition to matching for tongue tip-blade posture and location, sibilants must be identical: e.g. ziiz-a 'backbone' *ziis-a. In Ts'amakko, harmony operates from left to right, that is, progressively; the trigger is palatoalveolar and the target an alveolar.

In some languages such as Ineseño Chumash (Applegate 1972; Poser 1982), alveolars may trigger harmony. The rightmost sibilant determines the tongue tip-blade realization of all sibilants in the stem. In (3a), the 3sg. subject prefix is /s-/, but it is realized as [ʃ] if there is a palatal sibilant to its right in the stem (3b). In contrast, the dual marker /iʃ-/ (3c) is realized as [is] if followed by an alveolar sibilant (3d).

(3)	a.	/s-ixut/	[sixut]	'it burns'
	b.	/s-ilakʃ/	[ʃilakʃ]	'it is soft'
	c.	/p-iʃ-al-nan'/	[piʃanan']	'don't you two go'
	d.	/s-iʃ-tiʃi-yep-us/	[sistisiyepus]	'they two show him'

Dental harmony is found in Nilotic languages. It operates between dental and alveolar stops, including nasals if a contrast exists in the language, and it may be triggered by consonants at either place. In Päri (Andersen 1988; Hansson 2001b, 2010) dental harmony is respected in roots (4a). Root-final stops that are the product of final mutation combined with affixation match the dental or alveolar property of the initial stop (4b).

(4)	a.	n̪ɔt̪	'sucking'		
		àtwáːtˋ	'adult male elephant'		
	b.	dèːl	'skin'	dèːnd-á	'my skin'
		t̪ùol	'snake'	t̪ùon̪d̪-á	'my snake'

Retroflex harmony is reported for Gimira (Benchnon), an Omotic language of Ethiopia (Breeze 1990). In this language, coronal sibilants in roots match in retroflexion (and tongue tip-blade distinctions s/ʃ) and a causative suffix /-s/ agrees for the retroflex feature with a preceding root consonant across intervening vowels and consonants, including non-retroflex /r/. Numbers indicate tone levels.

(5)	a.	mak 2	'say'	mas 2	'cause to say'
		dub 4	'dance'	dus 4	'cause to dance'
	b.	z̪ert 1	'be red'	z̪erṣ 1	'make red'
		t̪ʂ'ud'	'spit'	t̪ʂ'uʂ'	'cause to spit'

Retroflex harmony can affect sibilants, as in Gimira, or operate between oral or nasal stops, as in Australian languages such as Arrernte (Arandic) (Henderson 1988; Tabain and Rickard 2007) in which apical alveolar and retroflex stops match for retroflexion in a root. Arsenault and Kochetov (to appear) report that Kalasha roots exhibit retroflex consonant harmony between stops, fricatives, and affricates, respectively, but only when participating consonants have the same manner of articulation. Root-internal combinations of retroflexes and non-retroflexes with the same manner of articulation are rare or unattested. If manner differs, retroflex, and non-retroflex consonants freely combine. In all the reported cases, retroflex harmony appears to be sensitive to manner distinctions.

2.1.2 Nasal Harmony In nasal consonant harmony, nasal stops typically harmonize with voiced stops and oral sonorant consonants. Nasal consonant harmony is attested primarily in Bantu languages. In Yaka (Hyman 1995), a nasal stop in a root causes a /d/ or /l/ in the perfective suffix (6a–c) to become [n] (6d–f). Prenasalized stops are not triggers (6c) and do not block harmony (6f). Vowel height harmony regulates the height quality of the suffix vowel.

(6)	a.	sól-ele	'deforest'	d.	kém-ene	'moan'
	b.	jád-idi	'spread'	e.	nútúk-ini	'bow'
	c.	kúːnd-idi	'bury'	f.	méːŋg-ene	'hate'

Intervening vowels and non-participating consonants are transparent to the harmony.

2.1.3 Liquid harmony Liquid harmony involves alternations between /r/ and /l/. In Bukusu (Bantu), it is attested in roots (Hansson 2001b, 2010). In addition, the benefactive suffix /-il-/ is realized as [-ir-] following a stem with [r] (Odden 1994a). Vowel height harmony causes the suffix vowel to lower to mid following mid vowels.

(7)	a.	teːx-el-a	'cook for'	d.	reːb-er-a	'ask for'
	b.	lim-il-a	'cultivate for'	e.	kar-ir-a	'twist'
	c.	iːl-il-a	'send thing'	f.	resj-er-a	'retrieve for'

In Sundanese (Malayo-Polynesian), /l/ triggers harmony of /r/ to [l] (Cohn 1992), as with the plural infix /-ar-/ in (8f).

(8)	a.	kusut	'messy'	d.	k-ar-usut	'messy' PL.
	b.	rahɨt	'wounded'	e.	r-ar-ahɨt	'wounded' PL.
	c.	ləga	'wide'	f.	l-al-əga	'wide'

2.1.4 Dorsal Harmony Dorsal harmony is found in Totonacan languages, and involves alternations between velar and uvular consonants. In Tlachichilco Tepehua (Watters 1988; Hansson 2001b, 2010), a uvular /q/ causes a preceding

velar to become uvular, which in turn conditions lowering of the preceding high vowel (9b).

(9)	a.	ʔuks-k'atsaː	[ʔuksk'atsaː]	'feel, experience sensation'
	b.	ʔuks-laqts'-in	[ʔoqslaqts'in]	'look at Y across surface'

In general, dorsal harmony targets velar consonants, altering them to uvular.

2.1.5 Laryngeal Harmony Laryngeal harmony requires consonants to agree for the laryngeal properties of aspiration, glottalic airstream, or voicing, as characterized by the features [spread glottis], [constricted glottis], and [voice], respectively. It appears frequently in morpheme structure constraints (MacEachern 1997 [1999]), but is rarer in patterns showing alternations.

Voicing and aspiration harmony is found in (non-click) stops in morphemes of Zulu (Bantu), as in (10a) (Khumalo 1987; Hansson 2001b, 2010). Loanwords (10b) also obey the restriction.

(10)	a.	ukú-peta	'to dig up'
		úku-pʰátʰa	'to hold'
		uku-guba	'to dig'
	b.	í-kʰôtʰo	'court'
		um-bídi	'conductor' < English 'beat'

Kera (Chadic) appears to have voicing alternations in affixes conditioned by voiced stops or affricates in the stem (Ebert 1979; Rose and Walker 2004), e.g. kə-sár-káŋ 'black (coll.)' vs. gə-ʤàr-gáŋ 'colorful (coll)'. However, Pearce (2005) has argued that voicing is actually conditioned by a neighboring low tone rather than the voiced stop in the stem. Hansson (2004b) also argues that in Yabem, a Huon Golf language of Papua New Guinea, voicing restrictions arose from tonal patterns rather than from consonant harmony.

Harmony for [constricted glottis] occurs in Chaha, a Semitic language of Ethiopia (Rose and Walker 2004), in which oral stops match for both [constricted glottis] and [voice]:

(11)	a.	jɨ-t'ək'ɨr	'he hides'
		jɨ-t'əβk'	'it is tight'
	b.	jɨ-kətf	'he hashes (meat)'
		jɨ-kəft	'he opens'
	c.	jɨ-gədɨr	'he puts to sleep'
		jɨ-dərg	'hits, fights'

Most exceptions to laryngeal harmony differ in both features (Rose and King 2007), ex. jɨ-gəmt' 'he chews off, gnaws' or jɨ-dəfk' 'he scrubs and pounds laundry'.

In addition to the main types reported here, (Hansson 2001b, 2007b) also lists stricture and secondary articulation harmonies. Stricture involves alternations

between stops and fricatives, and is attested in Yabem, e.g. se-dàgù? → [tédàgù?] 'they follow' REALIS. Secondary articulation refers to labialization, palatalization, velarization or pharyngealization. There are a few reported cases discussed in Hansson (2007b): pharyngealization in Tsilhqot'in (also known as Chilcotin, Athapaskan) (Cook 1976, 1983, 1993), which interacts with sibilant harmony, velarization in Pohnpeian (Micronesian) (Rehg and Sohl 1981; Mester 1988) and palatalization in Karaim (Turkic) (Kowalski 1929; Hamp 1976; Nevins and Vaux 2004a), as shown below:

(12) a. kʲumʲuʃʲ-sʲuzʲ 'penniless, unpaid'
b. kʰorkuv-suz 'fearless, safe'

In sum, consonant harmony targets a range of segments: dorsals, liquids, and coronal obstruents, as well as segments classified according to nasal and laryngeal features. A consistent characteristic is that the interacting segments share a high degree of similarity. Notably absent is harmony for major place features such as [labial], [coronal], or [dorsal], as well as classificatory features that tend not to assimilate even locally, such as [sonorant] or [consonantal].[2]

2.2 *Local Vowel-consonant Harmony*

Harmony in which contiguous strings of segments are affected is labeled *vowel-consonant harmony*. This type of local harmony involves vowels and consonants and can be triggered by either. Three main types are outlined: nasal harmony, emphasis harmony, and retroflex harmony.

2.2.1 Nasal Harmony Nasal harmony is triggered by nasal consonants or vowels, and affects vowels and certain consonants depending on the language. An example of nasal harmony triggered by vowels is found in Epena Pedee, a Choco language of Colombia (Harms 1985, 1994; Walker 1998[2000]):

(13) /peɾõɾa/ [peɾ̃õɾ̃ã] 'guagua (a groundhog-like animal)'
/ũbɯsi/ [ʔũᵐbɯsi] 'neck'
/wãhida/ [w̃ãh̃ĩⁿda] 'they went' (go PAST.PL.)
/wãĩtʰee/ [w̃ãiⁿtʰee] 'go' (future)
/dãwe/ [nãw̃ẽ] 'mother'
/kʰĩsia/ [kʰĩsiə][3] 'think'

In this language, nasal harmony is triggered by a nasal vowel and nasality spreads progressively onto vowels, glottals, glides, and liquids, but it is blocked by obstruents and the trill /r/.[4] Stops at the right edge of the harmonic domain are pre-nasalized. In addition, the onset of the syllable containing a nasal vowel is nasalized; in this position, voiced stops become fully nasal and fricatives are nasalized, but voiceless stops remain oral.

Nasal harmony triggered by nasal consonants is attested in Capanahua, a Panoan language (Loos 1969; Walker 1998 [2000]).

(14)	ʔõnãmpãn	'I will learn'	kajatãnaiʔ	'I went and jumped'
	põj̃ãn	'arm'	ʧipõŋki	'downriver'
	bĩmɯ	'fruit'	warãn	'squash'

In Capanahua, nasal stops trigger nasal harmony regressively to vowels, glides, and glottals, but harmony is blocked by obstruents and liquids.

Nasal harmony differs cross-linguistically in terms of which segments undergo nasalization and which block harmony, as documented in Walker (1998 [2000]). Prior foundational surveys on nasalization patterns include Schourup (1972), Piggott (1992), and Cohn (1993b, c) (note also Pulleyblank 1989). Cross-language variation occurs in a nested dependency relationship: vowels > glides > liquids > fricatives > stops. If liquids are nasalized, so will be more sonorous segments such as glides and vowels. If liquids block nasalization, so will less sonorous segments such as fricatives and stops. In the examples above, nasalization targets vowels, glides, and liquids in Epena Pedee, but obstruents block progressive harmony. In Capanahua, nasalization affects vowels and glides, whereas both obstruents and liquids block harmony. In Sundanese (Robins 1957; Cohn 1990), nasalization spreads progressively to vowels and laryngeals (15a), but is blocked by obstruents, liquids, and glides (15b):

(15)	a.	ɲãĩãn	'to wet'	b.	ŋãjak	'to sift'
		kumãh̃ã	'how?'		ŋũdag	'to pursue'
		mĩʔ̃ãsih	'to love'		mõlohok	'to stare'

In Applecross Scottish Gaelic (Celtic) (Ternes 1973), nasalization spreads from a stressed nasal vowel in the syllable and progressively to vowels, laryngeals, glides, liquids, and fricatives, but is blocked by obstruent stops:

(16)	/mãˈhar/	[mãˈh̃ãr̃]	'mother'
	/ʧ̃ĩanu/	[ʧ̃ĩãnũ]	'to do, make'
	/frĩaˈv/	[f̃r̃ĩãˈṽ]	'root' PL.
	/kʰɔ̃ispaxk/	[kʰɔ̃ĩs̃paxk]	'wasp'

Finally, in many South American languages, particularly of the Tucanoan family, voiceless stops and fricatives are transparent to nasal harmony, neither undergoing nasal harmony nor blocking it, as in Tuyuca (Barnes 1996). In this language, nasality spreads bidirectionally with no blocking, even by voiceless obstruents. Morphemes are oral or nasal; voiced stops are in complementary distribution with nasal stops in the harmonic domain.

(17)	wáa	'to go'	w̃ã́ã	'to illuminate'
	hoó	'banana'	h̃ṍ	'there'
	osó	'bat'	j̃õsó	'bird'
	bipí	'swollen'	mĩpí	'badger'
	sɨgé	'to follow'	tĩŋ́	'Yapara rapids'

Mòbà Yoruba (Benue-Congo) is also reported to have transparent voiced and voiceless obstruents (Ajíbóyè and Pulleyblank 2008; Piggott 2003; Archangeli and Pulleyblank 2007), e.g. /udũ/ → [ũdũ] 'lover of sweet things', /isĩ/ → [ĩsĩ] 'worship'.

2.2.2 Emphasis Harmony Emphasis harmony is a term for pharyngealization or uvularization harmony, sometimes labeled post-velar harmony (Shahin 2002), and is found in Arabic and Aramaic (Semitic) dialects, as well as Berber (Kossmann and Stroomer 1997). Emphasis in Arabic is normally treated as a process triggered by the emphatic coronal obstruents /tˤ dˤ sˤ ðˤ/ (or /zˤ/ in some dialects), which contrast with plain counterparts: e.g. dɛm 'blood' vs. dˤɛm 'he hugged' in Jordanian Arabic (Al-Masri and Jongman 2004). Moreover, other consonants such as liquids and labials may be emphatic, and pharyngeals can trigger emphasis harmony. Al-Ani (1970), Dolgopolsky (1977), Ghazeli (1977), Zawaydeh (1999), and Shahin (2002) provide evidence that emphasis harmony is uvularization, distinguished from pharyngealization. There is articulatory evidence for upper pharyngeal constriction, characteristic of uvulars. The primary acoustic effect of emphasis/uvularization is a large drop in F2 on neighboring vowels, while pharyngealization produces a high F1. Lehn (1963) and Watson (1999, 2002) note that in addition to tongue root retraction and general pharynx contraction, articulatory correlates may include labialization, lateral spreading, and concavity of the tongue. In Yemeni Arabic, labialization causes short high vowels to round in emphasis contexts: e.g. yimalliːhin 'he fills them F.' vs. yusˤaffiːhun 'he cleans them F.' (underlining indicates emphasis, here and below).

Examples of emphasis spread in Cairene Arabic are given below (Watson 1999: 274). See also Gairdner (1925), Harrell (1957), Lehn (1963), Abdel-Massih (1975) for discussion.

(18)	sˤubyaːn	'boys'	rabatˤ	'he bound'
	dˤarab	'he hit'	basˤala	'onion'
	maradˤ	'illness'	xaddˤar	'to make green'

Emphasis harmony minimally spreads to an adjacent vowel (Broselow 1976), a type of local assimilation, but maximally it extends across the entire phonological word, affecting both consonants and vowels. An example of word-level harmony in (19) is from Azerbaijani Jewish Aramaic (Hoberman 1988), where spreading is bidirectional. In this language, words are only rarely disharmonic, and affixes alternate in accordance with the harmonic root:

(19)	xarupa	'sharp'	xamusa	'sour'
	xarip-ula	'sharpness'	xamis-ula	'sourness'
	xoʃ	'good, pleasant'	razi	'satisfied'
	na-xoʃ	'ill, sick'	na-razi	'unsatisfied'

When emphasis harmony affects all segments in a word, it is sometimes difficult to identify the trigger, and this has led some researchers to propose that the

emphasis feature is a property of the syllable (Lehn 1963; Hoberman 1988) or a suprasegmental feature of the word (Harrell 1957; Tsereteli 1982) rather than a particular segment.

Emphasis harmony can be blocked by high (front) vowels and consonants. In Cairene Arabic, non-tautosyllabic high front vocoids optionally block harmony to the right of the emphatic consonant (20a); compare with (20b). It is not blocked by a tautosyllabic high front vowel (20c) or in leftward harmony (20d) (Watson 1999, 2002).

(20)	a.	sˤaːħib	'my friend M.'	b.	asˤħaːb	'friends M.'
		ʕasˤafiːr	'small birds'		asˤfuːr	'small bird'
	c.	taxdˤiːr	'making green'	d.	wisˤil	'he arrived'
		tˤifl	'child'		ʔamiːsˤ	'shirt'

In a southern Palestinian Arabic dialect discussed by Davis (1995a), emphasis harmony spreads bidirectionally. Leftward spreading is unimpeded (21a) but rightward spreading (21b) is blocked by high front segments /i j ʃ ʤ/ (21c). Similar effects are found in other dialects (Younes 1991; 1993).

(21)	a.	ʔabsaːtˤ	'happier'			
		naʃaːtˤ	'energy'			
		maʤassˤasˤiʃ	'it did not solidify'			
	b.	sˤabaːħ	'morning'	c.	sˤajjaːd	'fisher, hunter'
		ʔatˤfaːl	'children'		ʕatˤʃaːn	'thirsty'
		tˤuːb-ak	'your M.SG.blocks'		tˤiːn-ak	'your M.SG. clay'

Al-Masri and Jongman (2004) report that in Jordanian Arabic, high vowels in words such as /tˤubəh/ 'brick' exhibit lower F2 consistent with emphasis harmony, but this does not extend to the vowel in the next syllable. This contrasts with the pattern in words without high vowels. In Abu Shusha Palestinian Arabic (Shahin 2002), the obstruents /ʃ ʧ ʤ/ block emphasis harmony in both directions: regressive /ʕaʃarˤa/ → ʕaʃarˤa 'ten' or progressive /ʕatˤʃaːn/ → ʕatˤʃaːn 'thirsty' M.SG. but non-low vowels are transparent to harmony: /muhrˤ-aːt/ → mˤuhurˤaːtˤ 'fillies'.[5]

2.2.3 Retroflex Harmony Retroflex harmony in Sanskrit (Indo-Aryan) is a system which appears to operate at a distance between consonants. Retroflex continuants (/ʂ/ or /r/) cause a following dental nasal to become retroflex [ɳ], as shown for the nominal and adjectival suffix /-ana/ (Whitney 1889). This harmony applies freely over non-coronals and vowels (22a), but is blocked by other intervening consonantal coronals (22b):

(22)	a.	rakʂaɳa	'protection'	b.	varḍana	'increase'
		krpaɳa	'miserable'		rocana	'shining'
		ākramaɳa	'striding'		vrjana	'enclosure'
		kʂayaɳa	'habitable'		ceṣṭana	'stirring'

Whitney (1889) assumes that the tongue remains in the retroflex position until it encounters another consonant which requires a different position of the front and tip-blade of the tongue. This explains the dental coronals' blocking. This interpretation is adopted by Flemming (1995), Gafos (1996 [1999]), Ní Chiosáin and Padgett (1997), and others. Under this analysis, retroflex harmony constitutes vowel-consonant harmony rather than consonant harmony. To explain the fact that retroflex consonants also block harmony, Gafos (1996 [1999]) argues that only retroflex *continuants* can act as triggers, although Ní Chiosáin and Padgett (1997) argue against this position. Nevertheless, Hansson (2001b, 2010) argues that four typological properties of Sanskrit retroflex harmony set it apart from consonant harmony: opacity, disjoint sets of triggers and targets, directionality, and the domain of harmony (which may be phrasal).

Like Sanskrit, Kinyarwanda (Bantu) has a retroflex harmony that shows blocking. Harmony is regressive and causes an alveolar fricative to become retroflex when it precedes a retroflex fricative in a stem (Walker and Mpiranya 2006; Walker, Byrd, and Mpiranya 2008). Harmony is blocked by alveolar stops and affricates, retroflex stops, and palatal consonants. Intervening vowels and non-coronal consonants do not block harmony and are not perceptibly affected. However, an articulatory study of non-coronal consonants that intervene between harmonizing fricatives reveals that a retroflex posture is actually present during them (Walker *et al.* 2008). This finding is suggestive that the harmony causes a retroflexion feature to be present during the segments that separate audibly harmonizing fricatives, that is, it is a vowel-consonant harmony.[6]

In other patterns, harmony from retroflex consonants can explicitly target vowels across another consonant, as in Mpakwithi (Northern Paman) (Evans 1995), e.g. gwapɹa → [ŋgwaᵗfɻa] 'is eating'.

In sum, there are relatively few vowel-consonant harmony types, being restricted to articulations that are compatible with both vowels and consonants such as nasality, tongue root retraction and retroflexion. Another striking difference between vowel-consonant harmony and consonant harmony is that vowel-consonant harmony exhibits blocking effects, whereas consonant harmony does not. This is a characteristic shared by vowel harmony, as discussed in Section 2.4.

2.3 Non-local Vowel-consonant Harmony

Non-local vowel-consonant harmony is relatively rare, and differs from local harmony. Odden (1980) discusses reported cases of vowels palatalizing or causing velar shifts across a single, unaffected, intervening consonant. Three cases involving greater distances are discussed here.

Faucal or retraction harmony is attested in Salish languages such as Snchitsuʔumshtsn (also known as Coeur d'Alene) and Spokane-Kalispel-Flathead (Bessell 1998). In Snchitsuʔumshtsn there is both local retraction (backing and/or lowering) of vowels adjacent to a faucal consonant (uvular and pharyngeal), as well as non-local retraction. In the following examples, /i/ is retracted to [ɛ] and /ɛ/ to [ɑ].

(23)	*Local*	u-ts̲ɛq^{w}	‘it’s bright red’		
	Non-local	t^{s}ɛ̲ʃ-a̲lq^{w}	‘he is tall’	cf. t^{s}íʃ-t	‘it is long’
		nɛ̲ʔ-sa̲tt^{ʃ}-ɛ̲ʔqs-n	‘crank (on a car)’	cf. sɛ́tt^{ʃ}-nt^{s}	‘he twisted it’

Bessell (1998) argues from acoustic evidence that intervening consonants are unaffected by the retraction. Tigre (Semitic) /ə/ is reported lowered to [a] preceding ejectives and pharyngeals across other vowels and consonants (Palmer 1956; Odden 1980), but Rose (1996) found that other vowels were slightly lowered and retracted, suggesting a similar process to that of Snchitsuʔumshtsn.

Sibe (Tungusic) has progressive rounding vowel harmony and a vowel system that distinguishes two basic heights: /i y ɨ u ɛ ø a ɔ/ (Li 1996). Velar consonants in suffixes are realized as uvular if a non-high vowel appears in the preceding stem. This alternation takes place even across a high vowel.

(24)	*Diminutive*		*Non-self perceived past tense*	
	ɨldɨ(n)-kɨn	‘bright’	tɨ-xɨ	‘to sit’
	muxuli(n)-kɨn	‘round’	iɕi-xɨ	‘to be enough’
	sula-qɨn	‘loose’	tykɛ-χɨ	‘to watch’
	adʑi(g)-qɨn	‘small’	fɔndʑi-χɨ	‘to ask’

In Harari (Semitic), the rightmost coronal consonant (except /r/) in the 2F.SG. non-perfective is palatalized by the suffix -i. Palatalization may occur on more than one consonant (25b,e) and affects both roots and prefixes (Rose 2004).

(25)		2M.SG.	2F.SG.	
	a.	ti-kafti	ti-kafʧ-i	‘open’
	b.	ti-bitasi	ti-bitaʃi / ti-biʧaʃi	‘rip’
	c.	ti-k’admi	ti-k’adʒm-i	‘advance’
	d.	ti-dinak’i	ti-diɲak’-i	‘be surprised’
	e.	ti-fit’ani	ti-fiʧ’aɲ-i	‘hurry’
	f.	ti-sabri	ti-ʃabr-i	‘break’
	g.	ti-barri	ʧi-barr-i	‘fly’

These three non-local harmony cases bear some similarities to local interactions. Faucal harmony has its roots in local retraction of vowels, which is also regressive (Bessell 1998). Velars can become uvular adjacent to non-high back vowels in Chemehuevi (Uto-Aztecan) (Press 1980), Zimakani (Trans New Guinea) (Voorhoeve 1970), and Turkana (Nilotic) (Dimmendaal 1983), and adjacent to non-high vowels in Yakut (Turkic) (Krueger 1962; Nevins and Vaux 2004b). A local version of the same palatalization process as Harari is found in the related language Amharic (Semitic) (Leslau 1995). Faucal harmony appears to have an affinity with vowel-consonant harmony, but the other two cases share a stronger resemblance to consonant harmony. Dorsal harmony is described in Section 2.1.4 and Hansson (2001b, 2010) provides cases of palatal alternations that involve stops and fricatives. None of these cases exhibit blocking effects. Finally, the Sibe and Harari cases are triggered by or target specific affixes.

2.4 Vowel Harmony

Vowel harmony refers to assimilations among vowels. The assimilating vowels may be separated by consonants. Vowel harmony typically occurs within a word or smaller domain. Assimilations are observed for individual features and for feature clusters. We discuss examples of harmony for backness, rounding, height, and tongue root advancement/retraction, as well as harmony for all vowel place features. We use "vowel place" as a cover term to refer to features that are typically classified as vocalic, that is, ones applicable to backness, rounding, height, and tongue root posture.

There have been numerous valuable comparative studies of vowel harmony. Cumulatively, these have given shape to our understanding of the range of patterns across languages and the surrounding theoretical issues. Some examples of cross-linguistic generative studies cross-cutting several vowel harmony types include Kiparsky (1973c), Ringen (1975 [1988]), McCormick (1982), Goldsmith (1985), Cole (1987), Calabrese (1988, 1995, 2005), van der Hulst (1988b), Odden (1991), van der Hulst and van de Weijer (1995), Li (1996), Majors (1998), Polgárdi (1998), Krämer (2003), Nevins (2004), and Archangeli and Pulleyblank (2007), among many others. We point to some studies focused on harmony for a particular feature or feature group where relevant below.

2.4.1 Backness Harmony Tuvan (Turkic) shows a backness (palatal) harmony (Anderson and Harrison 1999; Harrison 1999, 2000). Tuvan presents eight vowel qualities, as shown in (26). Vowels may be long or short.

(26)

	Front	*Back*
High	i y	ɯ u
Non-high	e ø	a o

In native Tuvan words, vowels in a word are alike in backness, being drawn either from the set of front vowels (27a) or back vowels (27b).[7] The harmony produces alternations in suffix vowels, which take their cue from the backness of the preceding vowel, as shown in (27c). Within roots, there is reason to postulate that backness harmony is progressive. The evidence comes from epenthetic vowels in word-medial syllables of a disharmonic loan; these vowels harmonize with the vowel of the preceding syllable rather than the following one, for example, tex̲ina:r (from Russian tex'nar 'grain alcohol'), partu̲fel (from Russian part'fʲel 'wallet').

(27)

a.	ivi	'deer'
	idegel	'hope'
	xylymzyre:r	'smile' FUT.
	e:ren	'totem'
	xø:mej	'throat singing'

b.	ɯrak	‘far’	
	ulu	‘dragon’	
	ajɯːl	‘danger’	
	oruk	‘road’	
c.	is-ter-im-den	‘footprint’	PL-1-ABL.
	at-tar-ɯm-dan	‘name’	PL-1-ABL.
	esker-be-di-m	‘notice’	NEG-PST.II-1
	udu-va-dɯ-m	‘sleep’	NEG-PST.II-1

In this system, harmony is fully combinative: every vowel can be a trigger and a target for backness harmony. Many languages with backness harmony belong to the Ural-Altaic family. Well-examined cases include Turkish (e.g. Clements and Sezer 1982), Hungarian (Vago 1973, 1974; Ringen 1975 [1988]), and Finnish (e.g. Kiparsky 1973c; Ringen 1975 [1988]) (with numerous subsequent studies).

2.4.2 Round Harmony A cross-linguistic survey of round harmony is found in Kaun (1995, 2004). An example of round harmony occurs in the Halh (Khalkha) dialect of Mongolian (Mongolic) (Svantesson *et al.* 2005). Vowel phonemes of Halh are given in (28). The vowels on the right are characterized as pharyngeal. Their tongue root is pulled back and there is possible involvement of the pharyngeal constrictor muscles (Svantesson 1985; Svantesson *et al.* 2005). The pharyngeal characterization amounts roughly to the retracted tongue root feature specification [(+)RTR].[8] Vowel length is contrastive for all vowels, except that short [e] occurs only in non-initial syllables. In non-initial syllables, vowels are full or reduced.

(28)

	Non-pharyngeal		*Pharyngeal*	
	Unround	*Round*	*Unround*	*Round*
High	i	u		ʊ
Non-high	e	o	a	ɔ

Round harmony occurs among non-high vowels only, producing suffixal alternations between e/o and a/ɔ, as shown in (29a). A pharyngeal harmony is also seen in these data. High round vowels do not trigger round harmony and they block it from a preceding non-high round vowel (29b). However, /i/ is transparent. It may intervene between harmonizing non-high vowels while remaining unround itself (29c). If /i/ is the only stem vowel, the suffix is unround (29d). When round harmony is prevented, non-high vowels are unround in non-initial syllables.

(29)	a.	og-ɮo	‘to give’	DPST.
		xeːɮ-ɮe	‘to decorate’	DPST.
		ɔr-ɮɔ	‘to enter’	DPST.
		jaw-ɮa	‘to go’	DPST.
	b.	suːɮ-e	‘tail’	REFL.
		mʊːr-a	‘cat’	REFL.
		og-uɮ-ɮe	‘to give’	CAUS-DPST.
		ɔr-ʊɮ-ɮa	‘to enter’	CAUS-DPST.

c.	poːr-ig-o	'kidney' ACC-REFL.
	xɔːɧ-ig-ɔ	'food' ACC-REFL.
	ɔɧʲi-ɧɔ	'to squint' DPST.
d.	piːr-e	'brush' REFL.
	it-ɧe	'to eat' DPST.

It is noteworthy that Halh round harmony restricts triggers to non-high vowels. In addition, it shows an identity effect in that the harmonizing vowels must match in height (Kaun 1995, 2004).

Languages that show round harmony often also show harmony for another feature. A number of languages show both round harmony and backness harmony. Turkish (Turkic) is a well-studied case (e.g. Clements and Sezer 1982). Other examples include Tuvan (Harrison 2000), Tunica (Gulf) and Ewe (Kwa) (Odden 1991). Round harmony occurs together with pharyngeal (or [RTR]) harmony in several Mongolic languages (Svantesson *et al.* 2005) and with [RTR] harmony in certain Tungusic languages (Li 1996). Round harmony is reported to occur with ATR harmony in the Niger-Congo languages, Chumburung, Dagaare, and Igbo (van der Hulst and van de Weijer 1995; Krämer 2003), but as Krämer points out it is possible to analyze the harmonizing feature as [back] in place of [round].

2.4.3 Height Harmony Generative cross-linguistic studies dealing with vowel height harmony include Goad (1993) and Parkinson (1996). Kisa (Bantu) presents a harmony that causes lowering of high vowels (Sample 1976; Hyman 1998, 1999). Kisa's vowel inventory consists of [i e a o u]. Each vowel may be short or long. As shown in (30), the vowel in the suffix /-il/ lowers to mid when preceded by a mid vowel in the stem. The suffix remains high following a syllable with a high vowel or low [a].

(30) *Applicative*

-tsom-el-a	'pierce'
-rek-el-a	'set trap'
-βis-il-a	'hide'
-fuːng-il-a	'lock'
-βaːmb-il-a	'spread out, fasten down'

When the target vowel is /u/, height harmony occurs only if the trigger is /o/ not /e/, as shown in (31) with the reversive transitive suffix /-ul/. This rounding restriction is another example of an identity effect.

(31)

-tsom-ol-a	'pull out'
-rek-ul-a	'spring trap'
-βis-ul-a	'reveal'
-fuːng-ul-a	'unlock'
-βaːmb-ul-a	'spread apart, open up'

The Asturian dialect of Lena (Romance) shows a height harmony that involves raising (Hualde 1989a). The vowel inventory is /i e a o u/. In Lena, a high vowel in an inflectional suffix causes raising of a preceding non-high stressed vowel. The height harmony is stepwise: /e o/ raise to [i u] (32a) and /a/ raises to [e] (32b).[9] High stressed vowels are unaffected (32c).

(32)	a.	ka'biθu	ka'beθos	'head' M.SG./M.PL.
		fon'diru	fon'dera	'lower' M.SG./F.SG.
		re'undu	re'ondo	'round' M.SG./MASS.
		'tsubu	'tsobos	'wolf' M.SG./M.PL.
	b.	'eltu	'alta	'tall' M.SG./F.SG.
		ben'tenu	ben'tanos	'window' M.SG./M.PL.
	c.	ka'britu	ka'brita	'kid, young goat' M.SG./F.SG.
		'kubu	'kubos	'pail' M.SG./M.PL.

Lena's harmony demonstrates certain types of restrictions on triggers and targets. Targets must be stressed. Comparative studies in the generative tradition of height harmonies that affect stressed targets include Calabrese (1985), Hualde (1989a), Kaze (1989), Flemming (1993), Dyck (1995), and Walker (2005). Harmonies where a post-tonic high vowel causes raising of a stressed vowel are traditionally referred to as *metaphony* in Romance. Similar patterns involving fronting (and sometimes also raising) are referred to as *umlaut*, especially in Germanic. Triggers in Lena are subject to a morphological restriction: they must belong to an inflectional affix. This is evidenced by the form [sili'kotikos] 'suffering from silicosis' M.PL., where post-tonic /i/ in the stem does not trigger raising of the stressed vowel. Compare [sili'kutiku] M.SG., which shows that the stressed stem vowel does raise when the word contains an inflectional high vowel.

Harmony for the feature [low] is much less common cross-linguistically. Among patterns that have been treated as [low] harmony, controversy exists about the choice of feature. Relevant discussion can be found in van der Hulst (1988b), Clements (1991), Goad (1993), Archangeli and Pulleyblank (1994), Beckman (1995), van der Hulst and van de Weijer (1995), Casali (1996 [1998]), Parkinson (1996), Pulleyblank (1996), Hyman (1999), and Paster (2004), among others. For some harmony systems, there is debate about whether height features and [ATR] are both involved, as we discuss in Section 3.3.2.

2.4.4 Tongue Root Harmony The Pulaar dialect of Fula (Niger-Congo) shows a tongue root harmony (Paradis 1992; Archangeli and Pulleyblank 1994). Pulaar presents seven vowels, [+ATR] [i e o u] and [–ATR] [ɛ a ɔ].[10] In non-final position, the mid vowels' [ATR] specification is determined by the following vowel, producing alternations between e ~ ɛ and o ~ ɔ (33a). [ATR] harmony targets neither the high nor low vowels, causing potential disruptions in the harmonic sequence, but these vowels still trigger harmony in the preceding syllable (33b).

(33)	a.	ᵐbeːl-u	ᵐbɛːl-ɔn	‘shadow’ SG./DIM.PL.	
		peːc-i	pɛːc-ɔn	‘slit’ PL./DIM.PL.	
		dɔg-oː-ru	ⁿdɔg-ɔ-w-ɔn	‘runner’ SG./DIM.PL.	
		lef-el	lɛf-ɔn	‘ribbon’ CL.8.DIM.SG./CL.21.DIM.PL.	
		ceːlt-ol	cɛːlt-ɔn	‘cut’ CL.9/CL.21.DIM.PL.	
	b.	ɓet-ir-dɛ	‘weigh with’	ɓɛt-dɛ	‘weigh’
		ɗokk-iɗ-dɛ	‘become one-eyed’	ɗɔkk-ɔ	‘one-eyed person’
		feƒƒ-u-dɛ	‘fell’	fɛƒƒ-a	‘fell’ IPFV.
		bɔːt-aː-ri	‘lunch’	*boːt-aː-ri	
		ᵑgɔr-aː-gu	‘courage’	*ᵑgor-aː-gu	
		kɛlanᵑgel	‘break’ DIM	*kelanᵑgel	

Pulaar exemplifies a cross-linguistically common avoidance of [+ATR] low vowels and [–ATR] high vowels (Calabrese 1988; Archangeli and Pulleyblank 1994). As a result, tongue root harmony systems show a greater tendency to affect mid vowels.

Some tongue root harmonies exhibit what is known as a *dominant* pattern. An example is found in Maasai (Nilotic) (Tucker and Mpaayei 1955; Archangeli and Pulleyblank 1994; Baković 2000, 2002). Maasai has nine vowels [i ɪ e ɛ a ɔ o ʊ u]. In ATR harmony, affixes show alternations in conformity with the root’s ATR specification (34a) (roots are underlined). Some suffixes are invariably [+ATR], and these cause the root and prefixes to become [+ATR], as shown by alternations in (34b).

(34)	a.	/kɪ-n̲orr-ʊ/	kiŋorru	‘we shall love’
		/kɪ-ɪd̲ɪm-ʊ/	kɪdɪmʊ	‘we shall be able’
		/mɪ-kɪ-i̲tok̲i/	mikintoki	‘we shall not do again’
		/mɪ-kɪ-raŋ̲/	mɪkɪraŋ	‘let us not sing’
	b.	/ɪsʊj̲-ɪʃɔ-re/	isujiʃore	‘wash with something!’
		/ɪsʊj̲-ɪʃɔ/	ɪsʊjɪʃɔ	‘wash!/do the washing!’
		/a-rɔk̲-u/	aroku	‘I become black’
		/a-tV-rɔk̲-a/	atɔrɔka	‘I became black’

In this pattern [+ATR] is known as *dominant* and [–ATR] as *recessive*. Patterns such as this are often characterized as ones in which the presence of an underlying specification for the dominant feature causes all vowels in the word to become [+ATR] (e.g. Cole 1987). Maasai and other dominant [+ATR] systems show directional asymmetries with respect to blocking by the low vowel /a/. See Section 4.1 for discussion.

Generative cross-linguistic studies of tongue root harmony include Archangeli and Pulleyblank (1994), Pulleyblank (1996), Zhang (1996), and Baković (2000). Some tongue root harmonies are treated as involving privative or binary [ATR] and likewise others as involving [RTR]. Li (1996) argues RTR and ATR vowel systems show typological differences in harmony patterns as well as in inventories and their historical development. He maintains that canonical Tungusic vowel

harmony is an RTR system, whereas typical ATR systems are widespread in African languages.

2.4.5 Complete Harmony Some harmony systems show assimilation for all vowel quality features, often referred to as *vowel copy* harmony. As discussed by Steriade (1987a), many patterns of this kind are restricted to vowels separated by no more than a laryngeal segment. Kashaya (Pomoan) (Buckley 1994) exhibits a translaryngeal harmony that produces vowel identity in native morphemes, as shown in (34a). Supralaryngeal consonants block harmony (35b).

(35)	a.	mihi'la	'west'
		we'ʔej	'yonder'
		waʔali	'cane'
		so'hoj	'seal'
		hu'ʔul	'a while ago'
	b.	biʔdu	'acorn'
		hoja	'scoring sticks'
		k'aʔli	'between'
		ho'p^{h}une	'white-footed mouse'

In some systems, harmony that produces identical vowels operates across other segments. Transguttural harmony (Rose 1996) is attested in Jibbāli (Semitic) (Hayward, Hayward, and Al-Tabūki 1988) and Iraqw (Cushitic) (Mous 1993), and vowel copy harmony occurs across coronal consonants in certain morphological contexts in Pulaar (Paradis and Prunet 1989).

To conclude, vowel harmony occurs for all of the primary vowel place characteristics, affecting backness, rounding, height, and tongue root advancement/retraction. Some harmonies show assimilation for feature clusters. Assimilation for all vowel place features is attested, although it is prone to restrictions on the intervening consonants – a condition seen less frequently for vowel harmonies that involve only a subset of features. In contrast, consonant harmony involving major place features is not attested in adult language. Likewise, vowel-consonant harmony is restricted to properties that are compatible with both consonants and vowels, which can include characteristics of vowel place – realized as a secondary articulation on a consonant – but does not include major consonant place.

3 Analyses of Principal Aspects of Harmony Systems

3.1 Autosegmental Representations

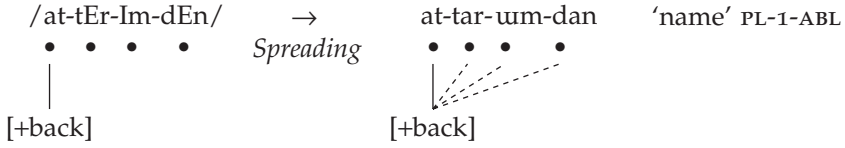
Autosegmental representations form the backbone of traditional non-linear analyses of harmony systems. We first review these accounts' primary features and then describe subsequent theoretical modifications and proposals. Recent progress points towards a greater diversity than previously conceived in the formal

motivations for harmony and the types of representations involved. These emerging differences motivate our discussions separately of advances in the analysis of (i) consonant harmony, (ii) vowel harmony and local vowel-consonant harmony, and (iii) non-local vowel-consonant harmony.

3.1.1 Spreading With the advent of non-linear phonology, the analysis of harmony systems as autosegmental spreading widely took hold (Goldsmith 1976a [1979]; Clements 1980). This marked a significant advance that directly or indirectly underlies many present-day treatments of particular harmony patterns.

In autosegmental representations, each harmonizing feature exists on its own tier. A feature's affiliation with a segment is represented by an association line which links the feature to the root node or to a node in the segment structure that the root dominates.[11] In autosegmental spreading, a feature links to additional segments, representing the fact that those segments have undergone assimilation for the feature. Spreading of [+back] from the root to suffix vowels in a Tuvan word is illustrated in (36). Bullets represent whatever node dominates [+back] in the segment structure. By convention, the broken line symbolizes a new association. For purposes of demonstration, we show nodes and linkage for vowels only here, returning to the status of consonants later. In the underlying form, we identify a suffix high vowel with *I* and a non-high vowel with *E*, indicating that they do not contribute an independent [±back] specification – that is determined by the vowel in the preceding syllable.

(36)

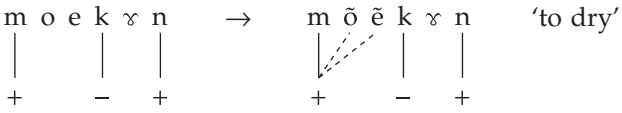


In autosegmental approaches to harmony, certain constraints or conditions on representations can be responsible for effects such as blocking and transparency, as we discuss below.

3.1.2 Targets, Blockers, and Transparent Segments Segments which block or are transparent to harmony are assumed to deviate from a canonical target. Targets can be identified through restrictions expressed in terms of autosegmental representations. For instance, targets could be restricted to segments that lack a pre-existing specification for the harmony feature, which is known as a *feature-filling* pattern. In approaches assuming an elaborated feature geometry, targets could be segments with the node that immediately dominates the spreading feature (e.g. Sagey 1986 [1991]; Steriade 1986; Archangeli and Pulleyblank 1987; Odden 1991), that is, they present a *target node*. Targets could additionally be subject to a requirement that they bear a certain feature specification. For example, round harmony in Turkish affects high vowels only; this could be analyzed with a restriction that the target be [+high] (or [–low]) (Goldsmith 1990).

Both vowel harmony and local vowel-consonant harmony are subject to blocking; blocking segments are described as *opaque*; we prefer the term *blocker* to minimize confusion with other cases of phonological opacity. The blocker segment halts harmony and in most cases does not undergo assimilation itself.

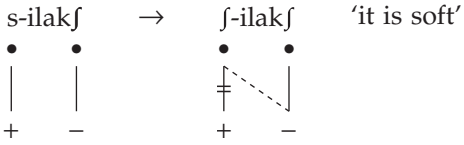
Early autosegmental accounts of blocking assumed that blockers are specified with the non-spreading value, for example [–nasal] if [+nasal] is spread. In their account of nasal harmony, van der Hulst and Smith (1982) propose that obstruents (or other blockers depending on the language) project a bound autosegment [–nasal]; [+nasal] associates to other segments via spreading until it encounters an association line, as shown for Sundanese in (37). The targets are the segments lacking nasal specification. The +/– specifications are shown here for the [nasal] feature tier.

(37) 

'to dry'

Essential to this treatment of blocking is the proposal that autosegmental representations are subject to the No-Crossing Constraint (NCC), which prohibits crossed association lines (Goldsmith 1976a [1979]). The NCC prevents a given feature from spreading over a feature specified on the same tier. A segment that is already specified for a feature on the tier on which spreading takes place can therefore block harmony (e.g. [k] in (37)), because spreading over that segment would cause line crossing. The NCC is widely assumed to be a "hard" universal, that is, it is never violated in linguistic structures.

Prespecification for the harmonizing feature does not always guarantee blocking. Some harmonies are characterized as *feature-changing*, which means that harmony can change a segment's specification for the assimilating feature. Sibilant harmony in Ineseño Chumash has been analyzed as feature-changing (Poser 1982, 1993; Lieber 1987; Shaw 1991), because they argue that both values of the relevant feature must be underlyingly specified. As discussed in Section 2.1.1, harmony causes /s/ to become [ʃ] and /ʃ/ to become [s], that is, when a sibilant is not followed by another sibilant in the word with which it must harmonize, its tongue tip-blade feature is contrastive and triggers harmony. An analysis of spreading [anterior] is illustrated in (38) (Shaw 1991).

(38) 

'it is soft'

In a feature-changing operation, the target acquires a feature specification and loses an existing one by removing its association line (usually followed by elimination

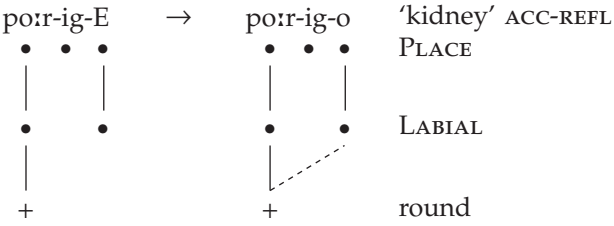
of the delinked feature). Delinking allows the NCC to be respected in the event that harmony persists onwards.

Whereas the absence of a specification for the harmonizing feature could cause a segment to be a target, it could also cause it to be transparent under an analysis that involves a *gapped* representation. A gapping analysis is illustrated in (39) for Halh round harmony. The spreading of [+round] skips transparent /i/ and reaches into the following suffix with a non-high vowel, that is, feature association gaps across /i/. (Assumptions vary as to whether association is regarded as gapping across the consonants here, as we discuss below.) A transparent vowel's failure to be targeted could be handled in various ways. For example, it could lack the target node or a feature specification that is a requirement for targets; alternatively, a markedness constraint could prevent formation of the vowel that would result if it were targeted.

(39) po:r-ig-E → po:r-ig-o 'kidney' ACC-REFL

A matter closely tied to transparency concerns locality restrictions on a feature's associations. Proposals have been made in which harmony phenomena can be parameterized according to the autosegmental tier or level of prosodic structure at which a trigger and target must be adjacent. For example, locality may be defined at the level of vowel nuclei or morae of adjacent syllables, allowing for transparent consonants in vowel harmony (Archangeli and Pulleyblank 1987, 1994). The theory of locality advanced by Odden (1994a) includes adjacency parameters that may require targets and triggers to belong to adjacent syllables or to have adjacent root nodes. Locality defined at tiers below the root in the feature geometry allows for transparent vowels in vowel harmony or transparent consonants in consonant harmony. Segments not specified on the relevant tier do not enter into the computation of locality. Transparent [i] in the Halh example above would thus lack particular feature specifications, as shown in (40) (the feature geometry shown is that of Sagey (1986 [1991]), but this is not essential). Here, the two mid vowels are "local" on the LABIAL tier.

(40) po:r-ig-E → po:r-ig-o 'kidney' ACC-REFL



Piggott (1996) proposes that harmony can take the form of a relation that holds at the suprasegmental levels of syllable or foot, and locality is defined in terms of

these categories (see also Piggott and van der Hulst 1997; Piggott 2003). Harmony must affect at least one segment within the syllable or foot domain, opening the possibility that certain other segments (e.g. obstruents in nasal vowel-consonant harmony) may be unaffected and hence transparent.

The well-formedness of gapped configurations has been questioned, spurring new analytical directions. Some work defines gapped configurations as ones where feature linkage skips over an eligible *anchor*, for example, a mora in vowel harmony (Archangeli and Pulleyblank 1994; Pulleyblank 1996). Under this interpretation, feature association in vowel harmony could skip over consonants in syllable margins but not vowels in syllable heads. Other analyses have taken the position that gapping may not occur across any segment (e.g. Ní Chiosáin and Padgett 1997, 2001; Walker 1998 [2000]; Rose and Walker 2004; Gafos 1996 [1999], 1998b makes a similar claim expressed in terms of articulatory gestures).[12] Approaches that make these locality assumptions must analyze (certain) transparency effects in ways other than the harmonizing feature skipping the transparent segment, as we discuss below.

3.2 *Analyses of Consonant Harmony*

Consonant harmony in generative phonology has also traditionally been analyzed using autosegmental spreading. Recent advances in the analysis of consonant harmony question this assumption and the underlying premise that harmony systems share particular characteristics (Hansson 2001b, 2010; Rose and Walker 2004). This is based on (i) a richer understanding of the range and typology of consonant harmony systems, (ii) a lack of blocking effects, and (iii) similarity between interacting consonants.

An autosegmental spreading analysis of consonant harmony must address the transparency of vowels and other consonants. Steriade (1987b) argued that intervening consonants and vowels in Ineseño Chumash sibilant harmony lack specification for the feature [anterior], a feature relevant only for coronals. This excludes vowels, dorsals, and labials from participation in the harmony, but the transparency of coronal consonants /t l n/ must be explained. This is achieved through underspecification. The sibilants contrast for [anterior] but /t l n/ do not have [–anterior] counterparts, and so are predictably [+anterior]. Predictable features are left unspecified. Harmonic spreading of the [anterior] feature operates unhindered between sibilants before a redundancy rule fills in predictable values on the other coronals. Shaw (1991) further argues that locality is defined on the [anterior] tier, so only segments specified for [anterior] (i.e. the sibilants) are involved in the locality calculation.

Shaw (1991) provides a taxonomy of consonant harmony systems which identifies two predominant systems: coronal harmony (=sibilant harmony) and laryngeal harmony.[13] This typology is in line with expectations concerning locality (or tier-based spreading) and underspecification. Features that are distinctively specified on consonants define an autosegmental tier not utilized by vowels, and these are precisely the features predicted to participate in consonant harmony. Segments

unspecified for such features should be transparent to harmony. Under a feature system in which vowels are specified as dorsal (Sagey 1986 [1991]; Steriade 1987a)[14] coronal consonants do not share features with vowels. Laryngeal features are used to distinguish among consonants but not vowels. Sonorants are inherently voiced, so they do not require specification of the feature [voice] (Itô and Mester 1986). The additional harmony types (dorsal, nasal, liquid) identified in Hansson (2001b, 2010) and Rose and Walker (2004) were not included in Shaw's typology, and were predicted not to occur. Yet, dorsal harmony would interfere with vowel specifications, and the features [nasal] and [lateral], being dependent on the root node in the feature geometry, were predicted not to spread across other segments.

Gafos (1996 [1999]) rejects tier-based locality, and argues that locality is defined in terms of articulatory gestures. Vowel gestures are contiguous across a consonant, whereas consonant gestures are not contiguous across a vowel (see also Ní Chiosáin and Padgett 1997, 2001). Given this version of locality, only coronal harmony, which involves assimilation for a tongue tip-blade feature, is predicted to be possible, as this is the only type of harmony which would not interfere with vowels. The tongue tip-blade is independent of the tongue dorsum used in the production of vowels, and its exact posture has no significant acoustic effect on vowel quality. Gafos proposes that tongue tip-blade features (Tongue Tip Constriction Orientation (TTCO) and Tongue Tip Constriction Area (TTCA)) do not skip over other segments, but are maintained through them with little perceptible effect. Coronal segments /t n l/ in Chumash harmony are predicted to alter their production in accordance with the harmonic domain in which they occur, either apical [n̺] in words like /k-sunon-us/ → [ksunonus] 'I obey him' or laminal [n̻] in words like /k-sunon-ʃ/ → [kʃunotʃ] 'I am obedient'. As stops do not contrast on this dimension in Chumash, they are not perceived as distinct. Other consonant harmony types (nasal, dorsal) are predicted not to occur, as they would involve interference with the tongue dorsum and other articulators.

However, Hansson (2001b, 2010) and Rose and Walker (2004) show that consonant harmonies are not restricted to coronals. Faced with a wider range of examples, both studies conclude that autosegmental spreading is inadequate as a model of consonant harmony. Consider nasal consonant harmony, as discussed in Section 2.1.2. Nasal consonants harmonize with voiced stops or sonorant consonants across other consonants and vowels. Yet, intervening vowels are unaffected by harmony and do not block, whereas in vowel-consonant nasal harmony, vowels are the prime targets of nasal harmony. They also can serve as triggers and some vowels may even block nasal harmony. Nasal consonant harmony does not behave as if autosegmental spreading of [nasal] is involved. Hansson (2001b, 2010) and Rose and Walker (2004) identify several key properties of consonant harmony which differentiate it from vowel-consonant harmony and vowel harmony. First, there are no blocking effects (although see Hansson 2007a). Second, the triggers and targets bear a high degree of similarity to one another. Third, Hansson (2001b, 2010) argues that there is no sensitivity to prosody in consonant harmony. Fourth, Hansson (2001a, b) argues that the predominant directional pattern in consonant harmony is right-to-left or regressive. There is no directionality tendency with

vowel-consonant harmony or vowel harmony (although see Hyman 2002). All of these factors point to an alternative perspective.

Hansson (2001a, b) and Rose and Walker (2004) propose that similarity is the driving factor in consonant harmony, and has its functional roots in speech production. Similar, but different, consonants present production difficulties that appear in speech errors (e.g. Fromkin 1971; Shattuck-Hufnagel and Klatt 1979; Frisch 1996). Sibilants are highly similar to one another and it is hypothesized that production is eased if they match for the position of the tongue tip-blade. Nasal stops harmonize with oral sonorants or voiced stops, which differ minimally from nasals. Voicing agreement occurs between obstruents, but is often restricted to stops. Homorganicity further contributes to similarity, and some laryngeal and nasal harmonies only operate between homorganic segments.[15] Rose and Walker (2004) determine similarity using the metric developed in Frisch, Pierrehumbert and Broe (2004). This metric assesses similarity based on shared natural classes of distinctive features in a given language by comparing the number of shared and unshared natural classes of two consonants. The size and contrastiveness of the segment inventory contributes to the similarity ratings. Natural classes, which incorporate the notion of contrastiveness, are better able to predict gradient phonotactics and capture major class subregularities than models based simply on distinctive feature specification. See also MacKenzie (2005) on the advantage of contrast-based similarity calculations.

Based on Walker (2000a, b), Hansson (2001b, 2010) and Rose and Walker (2004) develop an account of consonant harmony within Optimality Theory (OT), termed "agreement-by-correspondence," that establishes a correspondence relationship between similar segments, expressed as CORR-C↔C constraints, and indicated in the diagram below by coindexation. There is no autosegmental feature linkage between the segments.[16]

(41)

Cx	V	C	V	Cx
|				|
[αF]				[αF]

The CORR-C↔C constraints are arranged in a fixed implicational hierarchy from most similar to least similar. IDENTITY-CC constraints require the corresponding consonants to agree. Input-output faithfulness constraints can be placed between the CORR-C↔C constraints. The following tableau illustrates an example of nasal harmony in Kikongo (Bantu) for the word /futumuk-idi/ → [futumukini] 'resuscitated' (INTR) (Déreau 1955; Ao 1991).[17] The tableau shows only the stem and suffix portion of the word in which nasal harmony occurs. CORR-N↔D refers to homorganic nasal and voiced stop pairs, and CORR-N↔B to homorganic and heterorganic nasal-voiced stop pairs. Candidate (42b) has no correspondence relationship between /m/ and /d/, whereas candidate (42c) does, and the segments do not agree for nasality. Candidate (42a) has both a correspondence relationship and nasal agreement. It violates IDENT-OI(nas), which is violated by segments that gain a privative [nasal] specification in the output.

(42)

/futumuk-idi/	IDENT-CC (nas)	CORR-N↔D	CORR-N↔B	IDENT-OI (nas)
☞ a. futum$_x$ukin$_x$i				*
b. futum$_x$ukid$_y$i			*!	
c. futum$_x$ukid$_x$i	*!			

No correspondence relationship is established between the nasal and the voiceless stop /k/, as these two sounds are not sufficiently similar. Harmony in Kikongo operates between heterorganic segments, but if harmony were restricted to homorganic segments, the IDENT-OI constraint would occur between the homorganic CORR-N↔D and heterorganic CORR-N↔B constraints, causing candidate (42b) to win. Other work analyzing particular consonant harmony systems as involving corresponding segments or feature copy includes Clements (2001) and McCarthy (2007a).

The correspondence-based approach to consonant harmony allows similar consonants to agree at a distance; transparent segments are those that are not similar enough to participate in the harmony. No blocking is predicted, as lack of harmony is due to lack of/low ranking of correspondence with intervening segments.[18] This approach sets consonant harmony apart from vowel harmony and vowel-consonant harmony in using a different analytical mechanism.[19]

In conclusion, a more accurate typology of consonant harmony has led to alternate analytical devices using correspondence-based relations rather than autosegmental spreading. The assumption that all harmony systems are alike and therefore subject to the same type of analysis has been called into question, representing a significant departure in the analysis of consonant harmony and of harmony systems in general.

3.3 Analyses of Vowel Harmony and Local Vowel-Consonant Harmony

3.3.1 Harmony Imperative Whereas growth in our knowledge of the typology of consonant harmony points away from autosegmental spreading as a source for these systems, the situation is different for vowel harmony and local vowel-consonant harmony. In the majority, continued research on the latter harmony types supports representations for these systems in which a single occurrence of the harmonizing element is present throughout the sequence of segments (or anchors) that undergo assimilation. Most often this is modeled in terms of multiple linkage of an autosegmental feature. Some analyses based in the representations of Articulatory Phonology (Browman and Goldstein 1986, 1989, 1990) postulate gestures instead of features. Harmony is then analyzed as the temporal extension of a particular gesture over the interval that presents harmony (e.g. Gafos (1996 [1999]).

While analyses in rule-based phonology posit a spreading rule for the harmonizing feature, harmony is analyzed as driven by constraints in OT. Which type(s) of constraint are the cause of feature spreading is a topic of active debate. Among the issues at play are (i) whether harmony is driven by a spreading-specific constraint or is an epiphenomenon of independently motivated constraints, (ii) whether the constraint reflects a hypothesized phonetic basis for harmony, and (iii) the closeness-of-fit of predicted and attested patterns.

3.3.1.1 Spreading Constraint or Epiphenomenon? Vowel harmony and local vowel-consonant harmony have frequently been analyzed as driven by a constraint that requires features to align to a domain edge, such as the word, with spreading being the result. The feature alignment approach, first proposed by Kirchner (1993) for vowel harmony, is an extension of the Generalized Alignment constraint schema put forward by McCarthy and Prince (1993a). Applied to a backness harmony system, such as Tuvan's, the constraint would be ALIGN-R-([back], word), requiring that any [back] specification in a word be associated with the rightmost syllable (or segment) in the word. The constraint is interpreted as satisfied when the rightmost association of a [back] feature coincides with the rightmost syllable. Assuming that all vowels have a specification for backness, this constraint favors outputs in which all vowels are linked to a single specification for [back]. The basic analysis is illustrated in (43) with the Tuvan word [at-tar-ɯm-dan] 'name' PL–1-ABL.[20] For demonstration purposes, a hypothetical input is considered in which one of the suffixes is specified [–back].

(43)

/at-tEr-im-dEn/ | | + –	ALIGN-R([back], word)	IDENT-IO(back)
☞ a. at-tar-ɯm-dan | +		***
b. at-tar-im-den | | + –	*!*	**
c. at-ter-im-den | | + –	*!**	**

The winner in (43a) with full back harmony aligns the [+back] specification of the root to the right syllable of the word. A constraint not shown that enforces root faithfulness (e.g. McCarthy and Prince 1995; Beckman 1998 [1999]) is assumed to ensure preservation of the root's [back] specification. Further, a constraint on

locality is assumed to prevent feature linkage from skipping intervening syllables. Candidate (43b) spreads [+back] from the root to its neighboring suffix, and [–back] from the second suffix to the final syllable. This candidate violates the alignment constraint, because the [+back] specification is not right-aligned in the word. A gradient assessment is posited, such that a violation is accrued for each phonological unit by which a [back] specification is misaligned. Here, the unit is assumed to be the syllable, although in some analyses it could be the mora or segment. Candidate (43c) presents bidirectional spreading to neighboring suffixes from the second suffix. Like (43b), this candidate fails because it violates alignment. It accrues three violations because [+back] is misaligned by three syllables from the word's right edge.

Drawing on a proposal by Padgett (1995b), some analyses have formulated the spreading imperative as a constraint (SPREAD(F)) that requires a feature specification to be linked to all segments in a given domain. This approach does not stipulate directionality, allowing apparent directionality effects to follow from other properties of the system, such as positional faithfulness (e.g. Walker 2001b; Padgett 2002; see Kaun 1995 for a similar approach). Harmony patterns that show what is arguably true directionality has been analyzed with a directional version of a SPREAD constraint, as discussed in Section 4.1 below.

Also based in autosegmental representations is the *feature-driven markedness* analysis of harmony (Beckman 1997, 1998 [1999]), which seeks to subsume spreading-based patterns under constraints independently required for a variety of featural markedness effects. An essential claim is that violations of feature markedness constraints **F* are assessed at an autosegmental level. In other words, a constraint such as *[back] incurs a violation for each [back] specification in the representation, without reference to the number of segments with which each [back] specification is associated. Violations of *[back] can be minimized when a single feature is associated to all vowels in a word, thereby causing feature spreading.

The AGREE(F) approach departs from treatments of vowel harmony (and other assimilations) that intrinsically demand multiple associations of a single autosegment. The AGREE(F) constraint requires that adjacent elements have identical specifications for a feature without requiring that they share a feature specification (e.g. Lombardi 1999; Baković 2000).[21]

All of these approaches to the harmony imperative are typically non-specific about targets.[22] For example, alignment constraints cause feature spreading to seek a word edge rather than a particular target segment type. The lack of emphasis on targets is replaced by constraints that prevent certain segments from undergoing harmony. For example, the feature co-occurrence constraint *[+round, –high] prevents round harmony in Turkish from targeting non-high vowels (for discussion of this constraint, see Section 3.3.3).

This strategy differs from rule-based formalizations of the harmony imperative which permit positive target restrictions, for example requiring that targets be specified for a particular feature. In the case of Turkish round harmony, the target is required to be [+high] (see Section 3.1.2), thereby focusing on the segments that the assimilation affects. In the parametric rule formalism of Archangeli and Pulleyblank (1994), rules may include target conditions. These may be implicational

grounded path conditions, which govern well-formedness, for example, ATR/LO: 'if [+ATR], then [–low]'. These restrict targets to segments that obey the implication when they undergo the assimilation. Grounded conditions are similar in spirit to OT accounts in which markedness constraints prevent certain segments from undergoing harmony.

3.3.1.2 Phonetic Bases for Harmony A branch of research that has made significant advances in recent years centers on phonetic bases for harmony, and in some cases this has influenced the formalization of the harmony-driving constraint. Several studies have emphasized this topic. Hypothesized grounding or origins in phonetics fall broadly into two categories, perceptual and articulatory.

In the area of perceptual factors, certain vowel harmonies are identified as triggered by vowels with contrastive properties that have comparatively weak perceptual cues (e.g. Kaun 1995, 2004; Walker 2005b; see Suomi 1983 for related discussion). For example, Kaun argues that rounding is more perceptually subtle in non-high vowels than in high ones, causing non-high vowels alone to trigger round harmony in some patterns, such as Halh. The assumption that underlies this is that vowel harmony is primarily a perceptually driven phenomenon, an idea put forward by Suomi (1983). Harmony serves to increase the duration of the rounding feature, thereby enhancing the probability that it will be accurately perceived. Further, some harmonies are asymmetrically triggered by vowels in perceptually impoverished contexts (Ringen and Vago 1998; Walker 2005). For instance, metaphony-type harmonies are triggered by post-tonic vowels, often inflectional ones only (see Section 2.4.3). Harmony improves exposure of the harmonic feature either by extending it over a longer domain and/or by realizing it on a stressed vowel. In work on nasal vowel-consonant harmony, Sanders (2003) hypothesizes that certain patterns are motivated by an imperative to maximize distinctiveness between words for the perceptual dimension of nasality.

On articulatory bases, Majors (1998) notes that unstressed vowels undergo more vowel-to-vowel coarticulation than stressed ones, and she hypothesizes that patterns in which unstressed vowels assimilate to stressed ones have roots in coarticulation. Other research suggesting that some or all patterns of vowel harmony have origins in coarticulation includes Boyce (1988), Ohala (1994b), Steriade (1994), Manuel (1999), Beddor *et al.* (2001, 2002), Kaun (2004), and Linebaugh (2007). Boersma (1998, 2003) suggests that certain patterns of nasal harmony have an articulatory basis; in particular, they minimize the number of velum lowering and raising gestures.[23]

These issues surrounding hypothesized phonetic underpinnings and origins for harmony have been reflected with varying degrees of directness in the statement of harmony-related constraints. Versions of harmony-driving constraints have been proposed that express a restriction specifically over perceptually weak elements (e.g. Kaun 1995, 2004; Ringen and Vago 1998; Walker 2001b, 2005). Sanders' contrast-based analysis of nasal harmony utilizes constraints that explicitly require word pairs to have a certain degree of perceptual distance, which harmony reinforces. Boersma's functional analysis is implemented using constraints that penalize each movement of a particular articulator, such as the velum. Substantive considerations

have also given rise to position-sensitive constraints that may work in concert with other constraints to produce harmony patterns. Stressed syllables' resistance to change has been attributed to faithfulness constraints specific to prominent positions (Beckman 1998 [1999]). In addition, stressed syllables or other prominent positions are proposed to be the locus of markedness-based licensing, that is, a requirement that features have an association to a stressed position (e.g. Majors 1998; Walker 2004, 2005).

3.3.1.3 Attested Patterns and Constraints Another area of recent attention involves improving the harmony-driving constraint's closeness-of-fit with attested patterns. As discussed by Wilson (2002) and McCarthy (2004), problematic predictions regarding over- and undergeneration of patterns emerge with the primary formulations of the harmony imperative (e.g. ALIGN, SPREAD, *F, AGREE). A problem for AGREE is that it fails to capture harmony in forms which show partial harmony (e.g. because of blocking). It predicts instead that unless harmony is total, assimilation will fail altogether (see also McCarthy 2003b). Feature-driven markedness presents a similar problem.[24] Among the faulty predictions of ALIGN or SPREAD are the potential to (better) satisfy the harmony constraint by blocking epenthesis or deleting segments that are inaccessible to spreading because of blocking. Further, in the context of examining formal limits on possible constraints, McCarthy (2003b) challenges the gradient assessment of violations assumed for ALIGN and SPREAD constraints.

These issues have led to new proposals for the statement of the harmony imperative and/or the representations over which it operates. McCarthy (2004) proposes that harmony operates over *feature spans*. This approach circumvents problems of deletion or blocking of epenthesis because harmony is driven by avoidance of adjacent spans rather than a constraint that drives maximal spreading. Further, blocking segments initiate their own feature span, so segments intervening between a blocker and potential trigger can be compelled to undergo harmony as opposed to creating an independent span. A different solution to these issues is put forward by Wilson (2002, 2006c). He proposes to characterize the spreading constraint as *targeted*, which entails that the constraint identify both a marked structure and a repair.

Blumenfeld (2006) offers another take on certain harmony-driving constraints. A drawback has been noted for foot-bounded AGREE or markedness-based stressed-syllable licensing, applicable to patterns where a stressed vowel harmonizes with an unstressed one (e.g. Lena, Section 2.4.3). They make the unwanted prediction that the stress could shift to the unstressed syllable as a means of satisfying the constraint (Walker 2005; Blumenfeld 2006). To address this problem, Blumenfeld proposes *procedural* constraints, whose violation profiles are computed differently from standard constraints in OT in order to rule out certain processes. Procedural constraints are stated as implications, e.g. for foot-bounded AGREE, "if V_1 and V_2 are in the same foot, then they have the same value for [F]." A novel aspect of procedural constraints is that they cannot force the property in their antecedent to change. Thus, satisfaction of AGREE cannot be enforced by relocating foot boundaries, which could cause a shift in stress. The determination of where foot

boundaries would be located if AGREE were not present requires reference to the ranking of other constraints in the grammar.

Research on harmony imperatives thus has the potential to produce pivotal consequences for phonology theory. Proposals like those made by Wilson and Blumenfeld involve substantial departures from traditional constraint architecture in OT (Prince and Smolensky 2004). Further examination of the surrounding issues is needed to assess implications both for the analysis of harmony systems and the theory at large.

3.3.2 Feature Classes As mentioned in Section 2, some harmonies involve assimilation for more than one feature. Sets of features that pattern together frequently in phonological phenomena are referred to as feature classes, and they have spurred proposals to capture the recurrent co-patterning of particular feature clusters across languages. A significant approach is *feature geometry*, in which features are organized in a hierarchical structure in the segment (e.g. Clements 1985; Sagey 1986 [1991]; McCarthy 1988; Clements and Hume 1995, and many others). Nodes in the representation group features to form classes. For example, assimilations involving a feature class composed of [back] and [round] would involve spreading of the node that dominates these two features (e.g. Odden 1991), thereby characterizing combined back and round harmony as a unitary phenomenon.

However, the class node solution has limitations. Padgett (2002) points out that [back] and [round], which belong to a class he calls "color," show *partial class behavior* in Turkish vowel harmony. Whereas backness harmony targets all vowels, round harmony targets only high vowels. Thus, the Turkish genitive suffix /-In/, with a high vowel, has four alternants combining all rounding/backness combinations [-in -yn -ɯn -un], but the plural suffix /-lEr/ has only two alternants, front and back [-ler -lar], with vowels that are consistently unrounded. However, the class node approach predicts that either both [back] and [round] will spread or neither will spread in any given instance. Padgett dispenses with class nodes in the feature geometry and proposes a set-based notion of feature classes. For example, the set "Color" consists of the features [back] and [round], the set "Height" consists of [low] and [high], and so on. The sets, Color, Height, and so on, can be included in the statement of constraints or rules. Using this conception, Padgett employs constraint violability in OT to obtain partial class behavior. A constraint that requires spreading of Color features is dominated by a feature co-occurrence constraint *[+round, –high], as shown in (44) for the Turkish word [pul-lar] 'stamp' NOM-PL.

(44)

/pul-lEr/	*[+round, –high]	SPREAD(Color)
☞ a. pullar		*
b. pullor	*!	
c. puller		**!

Violations for SPREAD(Color) are shown for each feature in the Color set that fails to spread from the root. The winning candidate, (44a), shows harmony for [back] only, respecting the higher-ranked constraint against round non-high vowels. Candidate (44b) is ruled out by the dominating markedness constraint. Candidate (44c) shows harmony for neither [back] nor [round], incurring two violations of SPREAD(Color). A root-specific faithfulness constraint is assumed to protect round, non-high vowels that originate in the root. In the genitive singular form [pul-un], [round] may also spread, because it will not cause a violation of *[+round, –high].

In addition to handling partial class behavior, an advantage of the feature set approach is its ability to characterize feature classes that partially overlap with one another, because it is not constrained by geometry. For example, as mentioned in Section 2.4.2, round harmony occurs together with backness harmony in a number of languages and with tongue root (or pharyngeal) harmony in several others. This could be handled by postulating a Color feature class and a separate class composed of [round] and [RTR]. Standard models of feature geometry would not permit [round] to be organized under two class nodes with partially overlapping features. As a result, a geometric approach would unify only one of these paired harmonies.

A related issue in the topic of feature classes concerns reaching consensus on the features involved in a particular harmony system. For example, Odden (1991) argues that height harmony in Kimatuumbi (Bantu) involves assimilation for [ATR] and [high] features,[25] which he postulates form a class. Yet, different analyses without involving [ATR] have posited multiple occurrences of the same vowel height feature, corresponding to differences along an acoustic scale for F1 (Parkinson 1996) or hierarchically organized aperture features (Clements 1991).[26]

The vowel height models of Clements (1991) and Parkinson (1996) have also been applied to partial or stepwise height harmony systems. A different approach is proposed by Kirchner (1996), who employs a local conjunction (Smolensky 1993, 1997) of faithfulness constraints for the relevant harmonizing features. For example, as discussed in Section 2.4.3, Lena's height harmony causes the following raising effects /a/ → [e] and /e o/ → [i u]. At issue is preventing the raising of /a/ to [i]. Applying Kirchner's approach, the harmony-driving constraint would be dominated by a conjunction of the constraints IDENT-IO(high) and IDENT-IO(low). The conjunction is violated by any vowel that violates both of these constraints at once but not by vowels that violate only one of these constraints or that violate neither. The attested stepwise raising effects in Lena each involve violations of only one of the height faithfulness constraints. However, the unattested two-step raising of /a/ to [i] would violate the local conjunction, and is thus prevented.

In sum, OT has shed new light on certain aspects of harmonies that involve assimilation for clusters of features. The notion of constraint violability paired with a reduction in the complexity of feature geometry makes available the treatment of partial class behavior. Further, the additive effect of constraints through local conjunction finds utility in capturing partial assimilations. At the same time, debate persists on the precise set of features in some cases.

3.3.3 Blocking An important development in the treatment of blocking is the use of markedness constraints or well-formedness conditions. The insight is that a segment may block harmony when a marked segment would have resulted if it underwent harmony. Also significant is the increased emphasis on phonetic bases for segments' failure to participate in harmony. These issues received attention in Archangeli and Pulleyblank's (1994) Grounding Theory and have remained a focus of much work since.

In vowel harmony, ubiquitous cases of markedness-based blocking involve avoidance of high [–ATR] vowels and low [+ATR] vowels. For instance, as discussed in Section 2.4.4, in Pulaar, high vowels [i u] block [–ATR] assimilation and low [a] blocks [+ATR] assimilation. This has been analyzed using feature co-occurrence constraints *[+high, –ATR] and *[+low, +ATR] (for example, Krämer 2003; foundational work includes Calabrese 1988 and Archangeli and Pulleyblank 1994). In optimality-theoretic analyses, markedness constraints responsible for blocking dominate the harmony-driving constraint so that they may prevent its complete satisfaction. In Pulaar, the ranking *[+low, +ATR], *[+high, –ATR] >> Align-L([ATR], word) will select fully faithful [ⁿgɔraːgu] 'courage', with blocking of regressive [+ATR] assimilation by [a]. The markedness constraints rule out fully harmonic alternatives, that is, *[ⁿgoræːgu], where the low vowel becomes [+ATR], and *[ⁿgɔraːgʊ], where the high vowel becomes [–ATR] through progressive assimilation. Similarly, the markedness constraint *[+round, –high] has been employed for blocking of round harmony by non-high vowels (for example, Kirchner 1993; Kaun 1995) and blocking by round vowels in harmony for [–high] (Beckman 1997).

Parallel approaches are seen in the analysis of vowel-consonant harmony. Working in a rule-based framework, Davis (1995a) analyzes blocking of emphasis harmony by high front vowels using grounded path conditions, RTR/HI: If [RTR] then not [+high] and RTR/FR: If [RTR] then not [–back]. The former is roughly equivalent to the constraint *[+high, –ATR]. In some emphasis harmonies, blocking effects differ according to the direction of assimilation, as discussed in Section 4.1.

Blocking in nasal vowel-consonant harmony has been analyzed using nasal feature co-occurrence constraints. Walker (1998 [2000]) proposes a hierarchy of such constraints ranked according to the compatibility of [+nasal] with other features, with lowest compatibility ranked highest. Walker bases compatibility on factors of perception, articulation, and aerodynamics. Her analysis of Applecross Scottish Gaelic (see data in Section 2.2.1) is shown in (45), with the word kʰɔ̃ĩ̃spaxk 'wasp'. A rightward nasal spreading constraint dominates constraints that prohibit the co-occurrence of [+nasal] with fricatives, liquids, glides, vowels, and sonorant stops. (Walker's ranking and labels for these constraints are shown here.) A constraint prohibiting the co-occurrence of [+nasal] with obstruent stops is ranked over the spreading constraint, causing stops to block nasal harmony, as in candidate (45a). Candidate (45b), in which nasal harmony also spreads to obstruent stops, is prevented by the blocking constraint *NasObsStop.

(45)

/kʰɔ̃ispaxk/	*Nas ObsStop	Spread-R ([+nas], word)	*Nas Fricative	*Nas Liquid	*Nas Glide	*Nas Vowel	*Nas SonStop
☞ a. kʰɔ̃ĩs̃paxk		****	*			**	
b. kʰɔ̃ĩs̃p̃ãx̃k̃	*!*		**			***	

Walker analyzes cross-linguistic differences in blocking of nasal harmony as the effect of ranking the spreading constraint at different points in the nasal co-occurrence constraint hierarchy.[27] For example, for Capanahua, which nasalizes vowels and glides but not liquids or obstruents (see Section 2.2.1), the spreading constraint would be ranked between *NasLiquid and *NasGlide.

Contrast as a basis for blocking in harmony is emphasized in analyses of nasal harmony by Homer (1998) and Flemming (2004). In these accounts, nasal harmony is blocked when its occurrence would endanger or neutralize a perceptual contrast between nasal stops and certain other consonants in the language, as enforced by systemic constraints that explicitly require contrasts and that contrasts be perceptually distinct. A core idea is that contrast-centered constraints can stand in conflict with the constraint that drives nasal harmony. For example, Flemming proposes that flaps block nasal harmony in Johore Malay, while glides and vowels undergo it, because a nasalized flap [ɾ̃] is closer perceptually to a nasal stop than a nasalized glide or vowel. Since the weak contrast between the second consonant in hypothetical words [mãnã] and [mãɾ̃ã] is not tolerated in the language, harmony terminates when it reaches a flap, yielding [mãɾa]. Likewise, the blocking of retroflex harmony in Sanskrit by oral dental consonants (see Section 2.2.3) is attributed to contrast maintenance (Ní Chiosáin and Padgett 1997). In some OT approaches, faithfulness constraints carry the full responsibility of preserving contrast. Thus, related insights are captured by analyses that achieve blocking using faithfulness constraints for features in a given segment type. Proposals along these lines are made by Piggott (2003) for non-participant segments in nasal harmony and by Gafos (1996 [1999]) for certain blocking effects in Sanskrit's retroflex harmony.

Contrast has also played a prominent part in analyses couched in frameworks that are not purely constraint-driven. An approach with long-standing roots builds on the hypothesis that correlations exist between contrast, markedness, and feature underspecification. This analysis employs a system of contrast-sensitive feature specification, which, in the case of harmony phenomena, obtains the presence of features that block spreading or the absence of relevant structure in transparent segments (see Dresher, Piggott, and Rice 1994 for an overview of theories in which contrast affects feature specifications). A different proposal, assuming full specification, is put forward by Calabrese (1995), who permits certain rules and conditions to be sensitive to contrastive feature specifications only.

Nevertheless, analyses relying on contrast or markedness conditions cannot readily explain cases in which blockers undergo harmony. One case was mentioned in Section 2.2.2 with respect to Jordanian Arabic. High vowels which are themselves

affected by emphasis harmony still block harmony. Similarly, in Karajá, high [–ATR] vowels undergo regressive vowel harmony and become [+ATR] but block further spreading (Ribeiro 2003), e.g. kɔɗu-dĩ → [kɔɗuni] 'a type of turtle' *[koɗuni]. If the trigger is a mid vowel, blocking is optional. This case appears to require a distinction between underlying high vowels which trigger harmony and derived high vowels which act as blockers. Further research is needed on this issue.

Although the usual scenario in vowel harmony is for consonants to be transparent (see Section 3.3.4 for various viewpoints on this issue), there are nevertheless reported cases of consonants blocking harmony and/or initiating their own harmonic domain. The reverse scenario, in which vowels block consonant harmony is not attested; recall that lack of blocking effects of any kind is one of the motivations for an analysis of consonant harmony that does not employ feature spreading. In Turkish, palatalized/palatal consonants [lʲ gʲ kʲ] condition front vowels to their right: e.g. [usulʲ-y] 'method' ACC *usulʲ-u (Levi 2001). In models such as feature geometry or element phonology, this is explained by assuming that palatalized consonants have vocalic features, and interact with vowel-feature spreading. A similar assumption can be applied to glides in languages such as Bashkir (Turkic) (Poppe 1964; van der Hulst and van de Weijer 1995), although the behavior of glides appears to be language-specific. For example, in Turkish, glides do not block, and Levi (2001) argues for a non-vocalic representation of Turkish /j/.[28] Cases in which consonants participate in vowel harmony are problematic for syllable-based analyses of harmony, and appear to favor local segment-to-segment-based harmony.

Other types of cases of consonants affecting vowel harmony are attested. In Assamese (Indo-Aryan) (Mahanta 2007), nasal stops block regressive ATR vowel harmony. Paster (2004) discusses the case of Buchan Scots height harmony, which is blocked by intervening voiced obstruents or nasal-obstruent clusters. In some dialects of Italy, such as the dialect of Umbertide, vowel harmony among post-tonic vowels operates across liquids but not other consonants (Canalis 2009). In Nawuri (Kwa) (Casali 1995), labial consonants (not /w/) block rounding harmony, and in Warlpiri (Pama-Nyungan) (Nash 1979; van der Hulst and Smith 1985; Harvey and Baker 2005), labial consonants /p w/ block harmony that changes /u/ to [i]. Harvey and Baker (2005) assume that spreading applies locally to consonants as well, and that [–round] is blocked from associating to labial consonants; harmony is thus halted. In other round vowel harmony systems, labial consonants do not block, suggesting that a parameterization or constraint-ranking difference must be involved. All of these cases raise the issue of how "transparent" consonants really are, and which consonants have the potential to block, and for what reason. For example, labials are assumed to share features with round vowels, but blocking by labials does not occur in all round harmonies.

Finally, the number of consonants – or their prosodic position – rather than the consonants' quality can affect blocking. Codas block ATR vowel harmony in Lango (Nilotic) (Noonan 1992; Archangeli and Pulleyblank 1994)[29] and Assamese (Mahanta 2007). In Yucatec Maya (Mayan) (Krämer 2001), complete vowel copy is blocked by two consonants: e.g. [lub'-uk] 'fall' SUBJ. or [wen-ek] 'sleep' SUBJ. vs. [hèekn-ak] 'break' SUBJ *[hèekn-ek]. These cases are all analyzed as spreading

operating between vocalic morae. Locality is violated by skipping over a consonantal mora, and blocking results.

In vowel harmony systems, blocking is also seen to arise through identity effects. A well-established identity effect is seen in some patterns of round harmony, where assimilation is restricted to vowels of the same height, as in Halh (see Section 2.4.2). Identity conditions like these are known as *parasitic* harmony systems in the work of Cole and Trigo (1988). They propose a geometric explanation in which the trigger and target are required to share a particular contextual feature. Harmony is then restricted to the domain of the contextual feature.[30] This basic strategy for obtaining identity effects is also implemented by Cole and Kisseberth (1995b, c) in their analyses of harmony based on *feature domain* representations.

An insightful innovation on this topic has brought articulatory explanation to bear. For round harmony, Kaun (2004: 105) suggests that the height identity condition "reflects a phonetic imperative to avoid the need for articulatory adjustments in the execution of a single gesture." She proposes a *gestural uniformity* constraint for [round] that requires a multiply-linked [round] feature to have a uniform mechanism for its execution. This constraint will be violated when [round] is linked to vowels of different height, because a lip-rounding gesture is generally different in high versus non-high vowels (see also Kaun 1995).

Also related to articulation, Kaun (1995) has argued that a lower jaw position is antagonistic to lip rounding, giving rise to a constraint that penalizes round non-high vowels. This has been applied to round-parasitic height harmony restrictions that avoid generating non-high round vowels, for example, [o]. As discussed in Section 2.4.3, Kisa's height harmony causes /i/ to lower to [e] following a syllable with mid [e] or [o]. However, /u/ lowers to mid only following [o], that is, /o . . . u/ sequences become [o . . . o] but /e . . . u/ sequences remain unchanged. In her analysis of this pattern in the Bantu language, Shona, Beckman (1997) uses (the equivalent of) *[+round, –high] to cause blocking of [–high] harmony by /u/. However, [–high] harmony from [o] *can* produce lowering of /u/ to [o]. In Beckman's account, this is permitted because she postulates that both [+round] and [–high] are linked across [o . . . o] and she interprets the shared specifications as incurring a single violation with respect to *[+round, –high] (see discussion of feature-driven markedness in Section 3.3.1). A violation of *[+round, –high] will already exist for the trigger [o], so spreading in this sequence in particular will not produce additional violations. This could be conceptualized in gestural terms by reinterpreting *[+round, –high] as a constraint that penalizes the execution of lip rounding with a relatively low jaw position, without sensitivity to the dimension of its temporal extent. In other words, it is the articulatory configuration that is dispreferred without a difference in penalty for articulations of longer duration.

We note that while identity effects in vowel harmony are reminiscent of similarity conditions in consonant harmony (see Section 3.2), the hypothesized functional bases are distinct and the analyses have followed different paths in current theory.

In sum, blocking is a property of vowel harmony and vowel-consonant harmony and is generally attributed to markedness or contrast constraints on feature co-occurrence, or identity effects on the harmony system. Blocking of vowel harmony

by consonants is not common, and when it does occur can often be attributed to featural similarity between the vowel and consonant.

3.3.4 Transparency As with consonant harmony, transparency effects in vowel harmony and local vowel-consonant harmony is an area that has propelled new theoretical advances. Progress has been made both by research bringing new perspectives to already established data and by studies collecting new data observations.

A significant step forward came with a re-examination of what it means for a segment to be transparent. Various work has coalesced in support of a claim that some harmonies show *perceptual transparency* in which the assimilating property is actually present during so-called transparent segments, but without being perceived by listeners. The important consequence of this discovery is that it obviates the need to postulate skipping of transparent segments in these cases.

Instances where perceptual transparency is suggested are diverse. In the case of vowel harmony, Ní Chiosáin and Padgett (2001) argue that reported transparent consonants actually participate in the assimilation (see also Gafos 1996 [1999]). They claim that consonants may be perceived as transparent because the harmonizing feature has a low contrast potential in these segments. Further, experimental research on the articulation of reported transparent vowels in vowel harmonies of particular languages points to the harmonizing property being present during the vowel in question, although the effect might be sub-phonemic and not perceived by listeners (e.g. on Hungarian, Benus *et al.* 2004, Gafos and Benus 2006; Benus and Gafos 2007; on Kinande, Gick *et al.* 2006).[31] A similar analysis could be applied to the transparent non-low vowels in emphasis harmony (uvularization), as discussed by Shahin (2002). She analyzes these cases as phonological transparency (skipping of the vowel), as there is no steady-state lowered F2 in tokens with transparent vowels. Nevertheless, the vowels do exhibit F2 lowering at onset followed by a rise, unlike non-low vowels in non-emphatic contexts. Whether this is simply coarticulation or is enough to qualify as [RTR] extension through the vowel is not clear. Walker and Pullum (1999) suggest that glottal stops that are reported to be transparent in nasal harmony are actually participants. They point out that the velum is presumably lowered during glottal stops in these contexts, which they take to meet the criterion of being [+nasal]. However, the glottal closure prevents nasal airflow, resulting in the lack of a nasal percept (for discussion of surrounding issues see Cohn 1990, 1993b; Ní Chiosáin and Padgett 1997; Boersma 2003).

Retroflex harmonies in Kinyarwanda and Sanskrit are other cases of this type. As mentioned in Section 2.2.3, the harmonizing tip-blade feature or gesture is hypothesized to be present during reported transparent consonants and vowels, but without significant perceptual or contrastive results (Flemming 1995; Gafos 1996 [1999]; Ní Chiosáin and Padgett 1997; Hansson 2001b, 2010; Rose and Walker 2004; Walker and Mpiranya 2006). Indeed, evidence that the harmonizing gesture is present during non-coronal consonants that had previously been reported as transparent was found for Kinyarwanda (Walker *et al.* 2008).

Further arguments for consonants' participation in vowel harmony come from patterns in which a set of consonants are transparent and a different set act as blockers. Some examples where consonants block vowel harmony are discussed in Section 3.3.3. We consider here the Najdi dialect of Bedouin Arabic (Abboud 1979), discussed by McCarthy (1994a) and Gafos and Lombardi (1999). In non-final open syllables, short /a/ raises to a high vowel, as shown in (46a). However, /a/ does not raise when it is preceded by a guttural consonant or is followed by a sequence of a guttural consonant plus [a] (but not [u] or [i]) (46b). McCarthy attributes the non-raising to sharing of [pharyngeal] across the vowel(s) and guttural. Raising also does not occur when the /a/ is followed by an oral coronal sonorant plus [a] (46c). Like the guttural cases, this is analyzed as the vowel sharing a [pharyngeal] feature with the following /a/ and the intervening consonant.

(46)	a.	/katab/	[kɪ̲tab]	'he wrote'
		/nataf+aw/	[ntɪ̲faw]	'they (M) pulled feather'
		/kasar/	[kɪ̲sar]	'he broke'
		/sakan/	[sɪ̲kan]	'he dwelled'
		/ɟamal+uh/	[ɟmɪ̲luh]	'his camel'
		/samiʕ/	[sɪ̲miʕ]	'he heard'
		/ʃarib/	[ʃɪ̲rib]	'he drank'
	b.		[ʕa̲rif]	'he knew'
			[ʁa̲dar]	'he betrayed'
			[sa̲ʔal]	'he asked'
			[da̲χal]	'he entered'
		cf. /χadaʕ-uh/	[χdɪ̲ʕuh]	'he deceived him'
	c.		[ɟa̲las]	'he sat'
			[ɟa̲raf]	'he washed away'
			[ʃa̲nag]	'he beheaded'

The claim that the feature which causes the vowel to remain low is actually present in an intervening consonant is supported by the blocking of [pharyngeal] sharing by non-guttural obstruents in the first four examples in (46a). If these consonants are unable to undergo [pharyngeal] assimilation, as McCarthy suggests, raising should take place. On the treatment of coronal sonorants' receptiveness to participating in the harmony, see McCarthy (1994a) and Gafos and Lombardi (1999). Particularly relevant to our present concern is that the pattern points to a conclusion that "transparent" consonants participate in feature sharing/assimilation between vowels and that consonants block when they do not participate.

Despite many instances of purported skipping of segments being reduced to perceptual transparency in vowel harmony or local vowel-consonant harmony, the status of other cases remains to be investigated, and a residue exists for which this explanation does not appear promising. An example of the latter is transparent voiceless obstruents in nasal vowel-consonant harmony, for example, in Tuyuca (Section 2.2.1) and Guaraní (Tupí). An acoustic study of Guaraní's voiceless stops in nasal harmony contexts confirms that they are produced as voiceless and appear

to be oral (Walker 1999). Lena's height harmony presents another case. As described in Section 2.4.3, high inflectional vowels cause raising of preceding stressed /e a o/ to [i e u], respectively. In words with antepenultimate stress, a vowel in the penultimate syllable is transparent to harmony between the final and stressed vowels, as in (47) (Hualde 1989a).

(47)	a.	trw'ebanos	trw'ibanu	'beehive' (M PL./M SG.)
	b.	p'aʃara	p'eʃaru	'bird' (F SG./M SG.)

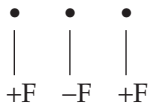
Although this phenomenon awaits instrumental investigation, it appears likely that unstressed /a/ genuinely does not undergo the height harmony (Walker 2004). It is improbable that harmony-induced raising of unstressed /a/ to [e] is not perceived, as [e] is an attested unstressed vowel quality in Lena.

The resulting theoretical picture is one in which transparency effects are represented in more than one way correlating with (at least) two types of distinct phenomena. In cases of perceptual transparency, segments actually undergo harmony, that is, they become specified for the harmonizing feature. In other patterns, transparent segments genuinely do not present the harmonizing feature. In many of cases of *genuine transparency*, the transparent segment could be prevented from presenting the harmonizing feature because of a phonetically-based markedness constraint. This seems tenable, for instance, in the case of transparent voiceless obstruents in nasal harmony, but seems less probable for Lena.

Harmony patterns that show genuine transparency have sparked numerous proposals, and apart from widespread agreement that these segments do not bear the harmonizing feature or gesture, there is little consensus on the particulars of their analysis. We examine some of the major concepts here.

Several proposals have emerged that preserve the claim that feature spreading may not skip segments. Several of these postulate separate but identical occurrences of the harmonizing feature specification before and after a transparent segment, as illustrated in (48), rather than a gapped feature linkage.

(48)



For example, in work on vowel harmony, Pulleyblank (1996) interprets violations of feature alignment in terms of what he calls *local domains*, which can drive identical feature specifications to flank a transparent segment. Nevins (2004) formalizes harmony as a feature copy procedure which searches out targets according to specific parameters. Walker (2004) proposes that stress-targeted harmonies (e.g. in Lena) operate over feature *chains*, which permits harmony to be achieved by a corresponding feature specification in a target syllable. Krämer (2003) characterizes transparency as *balance*, where a transparent vowel is required to either agree for the harmonizing feature with the vowels in both of its flanking syllables

or to disagree with both. Baković (2000) and Baković and Wilson (2000) use a targeted markedness constraint that disfavors forms in which certain segments undergo harmony. This constraint imposes a harmonic ordering on certain candidate pairs and interacts with AGREE to yield transparency.

In the area of nasal harmony, Sanders (2003) proposes that harmony in Tuyuca is driven by constraints on contrast, not spreading. Constraints on perceptual distinctness are better satisfied the more two words differ in nasality. However, a highly ranked markedness constraint on nasalized voiceless obstruents prevents these segments from being formed, at the cost of maximizing contrast; otherwise all segments in the word agree in nasality. Walker (1998 [2000]) analyzes transparency in nasal harmony as a kind of derivational opacity implemented in OT using sympathy theory (McCarthy 1999b). Her account builds on rule-based scenarios where a transparent segment undergoes spreading for the harmonizing feature, with the segment's harmonizing feature specification being subsequently altered, as driven, for example, by a markedness constraint (e.g. Clements 1980; Vago 1976; Piggott 1988b).

The above analyses prevent the need for gapped configurations, which permits more concrete representations where interruptions in an articulatory posture for a transparent segment are directly reflected in features' domains of association (or in gestures' extent of duration). However, other approaches allow gapped configurations or something similar. As mentioned in Section 3.1.2, gapping across intervening segments is a traditional approach to transparency in autosegmental spreading accounts. Some recent analyses of vowel harmony that employ this strategy, with locality defined other than by strict root adjacency, include Halle, Vaux, and Wolfe (2000), Uffmann (2004), and Calabrese (2005). See also Shahin (2002). Boersma (2003) proposes that transparent segments in nasal harmony in languages like Tuyuca cause violations of the Line Crossing Constraint, a structure that he couches in the context of the perceptual representations that he posits. In work making use of feature-based domains, the domain of a feature specification spans a continuous sequence of segments but it may fail to be realized on certain segments within its domain, in order to satisfy a higher-ranked constraint (Smolensky 1993; Cole and Kisseberth 1995a, b, c).

In sum, work tackling transparency has yielded new and recent discoveries, including experimental evidence that certain transparent segments are participants in harmony to some degree. Overall, while progress has been made on the analysis of vowel harmony and local vowel-consonant harmony, certain issues surrounding the harmony imperative, feature classes, blocking, and transparency all continue to be the topic of active investigation and debate.

3.4 Analyses of Non-local Vowel-consonant Harmony

Non-local vowel-consonant harmony poses problems for both autosegmental spreading and analyses that do not specify targets. There could be an expectation that they should display the same properties as local vowel-consonant harmonies, but in fact they do not. The transparent intervening segments that are responsible

for the non-local property do not fall neatly into the typology of transparent segments just identified. Due perhaps to the typological rarity of non-local vowel-consonant harmony, but also to the idiosyncrasies of these systems, no unified analysis has been forthcoming.

In Salish retraction, harmony from faucal consonants targets vowels, but skips consonants. Bessell (1998) assessed whether long-distance anticipatory coarticulation could have given rise to faucal harmony. However, coronals would then be expected to show some evidence of participation, as they have an antagonistic articulation with tongue root retraction and retracted pronunciations of coronals have been recorded in related languages with local coarticulation, such as Nxaʔmxcin. Instead, Bessell suggests that Salish non-local harmony arose from the strong compatibility of [RTR] with vowels, and its marked combination with consonants.

In Sibe and Harari, features associated with vowels target consonants, but skip certain intervening vowels and consonants. Recall that in Sibe, a velar consonant in a suffix becomes uvular if there is a preceding non-high vowel in the root. Li (1996: 307) proposes an analysis of Sibe whereby a redundant secondary vowel feature (A^2 in dependency phonology, equivalent to [RTR]) is spread from the non-high vowel to the consonant. Intervening high vowels, which are transparent, have primary vowel features, and are invisible to spreading by secondary features due to tier-based segregation. Nevins (2004) and Nevins and Vaux (2004b) present a different analysis of Sibe in which the feature [–high] spreads from non-high vowels to a suffix velar consonant. Since both [+ATR] and [–ATR] vowels can trigger the harmony, they argue that a height feature is responsible rather than a tongue root feature. Li (1996) notes that there is no [ATR]/[RTR] contrast for vowels in Sibe. Nevins and Vaux analyze transparency using contrastive visibility (Calabrese 2005) in which high vowels are unmarked and non-high vowels are marked in Sibe. The harmony rule specifies that only the marked value of the harmonic feature is "visible," namely [–high]. This analysis requires a specific target and theory of markedness-based spreading which calls into question other accounts of transparent segments. It remains to be seen how applicable this approach could be to other cases of non-local interaction.

In Harari, palatalization affects alveolar consonants at a distance from the trigger vowel, skipping over all non-targets including palato-alveolars and high and front vowels. If feature spreading were the mechanism by which harmony were achieved, blocking by segments specified for the spreading feature such as these would be expected. Rose (2004) proposes a correspondence-based agreement requirement between suffix and stem, and further parameterizes it to refer to specific targets: obstruents are favored over sonorants. Not using spreading avoids transparency problems, but the trigger and target are specifically singled out.

While local vowel-consonant analyses converge on spreading or extension of features/gestures, with attendant predicted blocking and transparency effects, there appears to be no unified analysis of non-local vowel-consonant harmony. The cases are sporadic and each presents unique properties. All of them probably developed in some manner from local coarticulations or assimilations that have become extended and/or morphologized.

4 Further Aspects of Harmony Systems

4.1 Directionality

Harmony can operate in a leftward (regressive) or rightward (progressive) direction, or bidirectionally. While many analyses incorporate directionality into rules or constraints, directionality has been argued to follow from morphological structure (Baković 2000; 2003). There have also been proposals of directionality bias, suggesting that the default direction for vowel harmony and consonant harmony is regressive (Hyman 2002; Hansson 2001a, b, 2010 respectively), connecting this to speech planning or other functional explanations.

In Yoruba, tongue root harmony is regressive from roots to prefixes (Archangeli and Pulleyblank 1989).

(49) a. O + gE + de [ogede] 'incantation'
b. O + gE + dɛ [ɔgɛdɛ] 'banana, plantain'

As mentioned in Section 3.3.1, in OT, a widespread approach to harmony uses directional alignment constraints (Kirchner 1993), which align the harmonic feature with edges of morphological domains, such as roots or words, or prosodic domains, such as feet or prosodic words. Regressive harmony in Yoruba would be expressed using Align-L([RTR], Word) (Pulleyblank 1996).

Some researchers have eschewed stipulating directionality. Clements (1981) proposed that unspecified segments trigger feature spreading from a specified segment due to a well-formedness requirement. In the Yoruba example above, spreading is automatically regressive to fill in [ATR] specifications on the prefixes. Other researchers have achieved directional effects through positional faithfulness (e.g. Kaun 1995; Beckman 1997; Walker 2001b), by requiring that strong positions (root-initial, stressed) preserve their features. Baković (2000, 2003) argues that the morphological affiliation of the segments is responsible, and that harmony is stem-controlled, operating from the root outwards to affixes. Yoruba only appears to have regressive harmony because it is a prefixing-only language. In a suffixing-only language like Tangale (Chadic) (Kidda 1985), harmony is progressive. In a language that combines prefixes and suffixes such as Akan (Tano) (Schachter and Fromkin 1968; Clements 1981), harmony is bidirectional. Stem-controlled harmony is expressed as a cyclic system, operating in successively larger domains. This is expressed formally by a faithfulness constraint Stem-Affixed form Identity (SA-Ident), requiring a stem in an affixed form to be identical to the unaffixed stem for a given feature.

In dominant-recessive systems, however, suffixes can cause roots to harmonize, as was shown for Maasai in Section 2.4.4. Such systems are bidirectional, with harmony operating from wherever the dominant [ATR] feature is located. Dominant-recessive systems are analyzed with SA-Ident outranked by constraints forcing harmony for the dominant feature (Baković 2000).

Dominant-recessive systems can display asymmetric directionality behavior with respect to vowels with no harmonic counterpart. In Maasai (Tucker and Mpaayei 1955), the low vowel /a/ has no [+ATR] counterpart. Such vowels may be opaque, transparent, or undergo *repairing* (Baković 2000), alternating with a vowel that is normally another vowel's counterpart. The vowel [a] occurs in [−ATR] words (50a). Progressive harmony from the root *repairs* /a/ by raising and rounding it to [o] (50b), but in regressive harmony triggered by a suffix (50c) /a/ is a blocker. It also fails to undergo harmony when a prefix (50d).

(50)	a.	/ɪn-lɪpɔŋ-a/	ɪlɪpɔŋa	'full-grown female'
	b.	/ɪn-mudoŋ-a/	imudoŋo	'kinship'
	c.	/ɛ-ɪpʊt-a-rɪ-ie/	ɛɪpʊtariyie	'it will get filled up'
	d.	/a-duŋ-akɪn-ie/	aduŋokinie	'I cut for s.o. with s.t.'

Archangeli and Pulleyblank (1994) express the directional asymmetry by imposing a grounded condition on the regressive rule, preventing the combination of [+ATR] and low: LO/ATR; there is no such condition on the progressive rule, and harmony applies to /a/.

Baković (2002) argues that directional stipulation predicts vowel harmony systems which do not occur, such as the opposite of Maasai: progressive blocking with regressive harmony.[32] A stem-control analysis predicts that systems may have blocking in both directions, or only in the regressive direction, as in Maasai. His analysis appeals to SA-IDENT for the non-harmonizing features affected by repairing: [low] and [round], protecting the stem from harmony in the regressive direction. Prefixal /a/ in Maasai is predicted to harmonize with the stem, but in fact it does not (50d). This follows from a directional analysis, but the stem-control analysis must treat prefixes as outside the harmonic domain or subject to special faithfulness.

There are vowel harmony systems that show no effects of stem control. Ribeiro (2003) presents data from Karajá, a Macro-Jê language spoken in Brazil, which has both prefixes and suffixes and a regressive [+ATR] dominant harmony. Harmony is triggered by affixes and clitics (51a, b), or by roots. Examples (51c, d) show a disharmonic root /ɗuhɔ/ in which the initial [+ATR] vowel /u/ triggers regressive harmony in prefixes. The second root vowel can become [+ATR] when followed by [+ATR] suffixes or clitics (51d).

(51)	a.	∅-r-a-kɔhɔɗɛ=r-e	[rakoho'ɗere]	'he/she hit'
	b.	bɛdɔ-dĩ	[bedo'ni]	'a type of *filhote*'
	c.	∅-r-ɔ-ɗuhɔ=rɛrɪ	[roʧu'hɔrɛrɪ]	'he is cursing'
	d.	∅-r-ɔ-ɗuhɔ=r-e	[roʧu'hore]	'he cursed'

This case presents a problem for proposals to reduce directionality to stem control, as examples such as (51c) show. The clitic /rɛrɪ/ is unaffected by harmony from the root. Similarly, Sasa (2004) argues that regressive directionality in Pulaar's ATR harmony cannot be reduced to effects of cyclicity or positional faithfulness.

Mahanta (2007) shows that Assamese also involves regressive ATR harmony and argues for a sequential markedness account, notably *[–ATR][+ATR]. Hyman (2002) speculates that a tendency for regressive directionality in vowel harmony in the absence of root control may be connected with the greater robustness of anticipatory vowel-to-vowel coarticulation. Nevertheless, cases of genuine directionality are not all regressive. Harrison (2000) finds evidence in Tuvan that backness harmony which affects epenthetic vowels is progressive in word-medial contexts (see Section 2.4.1).

Directionality in vowel-consonant and consonant harmony is not always reducible to stem control. In nasal vowel-consonant harmony, direction of spreading frequently has to be stipulated, and often occurs within roots. In Capanahua, nasal harmony is regressive, whereas in other languages with the same set of targets, such as Malay, it is progressive. Walker's (1998 [2000]) analysis of nasal harmony incorporates directionality into spreading constraints (Spread-R or Spread-L) to reflect this. Hansson (2001b, 2010) argues that cases of progressive consonant harmony can be described as stem-controlled, but regressive harmonies cannot. In Ineseño Chumash, suffixes trigger changes on roots. Hansson (2001a, b) relates the regressive bias to speech production. In speech production studies, anticipatory errors and assimilations are more common than perseverative (Dell, Burger, and Svec 1997). This is modeled in a serial-order theory of speech production in which look-ahead activation of a consonant being planned can cause an earlier segment – especially a similar one with shared activation – to anticipate its production.

Directionality may affect the extent of harmony. In Nawuri (Casali 2002; Hyman 2002), phrasal vowel harmony is unbounded in the regressive direction, but only affects a single high vowel in the progressive direction. Similarly, progressive emphasis harmony in Northern Palestinian Arabic (Davis 1995a) is limited to adjacent syllables. Furthermore, progressive emphasis harmony is subject to blocking, whereas regressive emphasis harmony is generally unrestricted. Davis (1995a) uses this to argue for "process-specific" spreading rules, a progressive rule with a grounded condition RTR/HI and RTR/FR, which prevents [RTR] combining with high or front segments, and a regressive rule with no target conditions. McCarthy (1997) instead achieves the directional effect through ranking, with directional harmony constraints: Align-RTR-Left >> RTR/HI&FR >> Align-RTR-Right >> Ident-ATR. Regressive RTR harmony is more important than respecting the markedness constraint. Watson (1999) suggests that regressive is the unmarked directionality for RTR harmony, and this is why it overrides grounded or markedness conditions. Greater restrictions are placed on the marked direction, limiting its application.

Different segments may be targets or triggers in different directions. In Kinande (Mutaka 1995; Archangeli and Pulleyblank 2002), regressive ATR vowel harmony targets all seven underlying vowels /i ɪ u ʊ ɛ ɔ a/. The low vowel /a/ is an undergoer or transparent (see Gick *et al.* 2006 on transparent low vowels in Kinande). In contrast, progressive harmony only operates between high vowels, and /a/ is opaque to harmony. Archangeli and Pulleyblank (2002) develop an

analysis of the directional asymmetry by utilizing only an ALIGN-L constraint. Progressive harmony follows not from ALIGN-R but from the grounded condition HI/ATR (if [+high], then [+ATR]) applying in different morphological domains.

In conclusion, recent research has put forward the hypothesis that there is a regressive directionality bias for vowel harmony, consonant harmony and some cases of vowel-consonant harmony. The source of this bias and its status in linguistic theory is still being explored. On the other hand, stem control may override the directionality bias in certain cases. This is clearly an area for future research.

4.2 Domains

Harmony can be delimited by its domain of application, referring to the maximal constituent to which harmony is confined. Although there was a general recognition of proximity requirements in prior work on harmony, these requirements have been formalized using phonological constituents such as syllable and foot. Perhaps the strongest recent advancement has been the development of licensing analyses of stress-based harmonic systems.

4.2.1 Phonological Domains Vowel-consonant harmony, both nasal and emphasis harmony, can be restricted to apply within the syllable. However, this can also be analyzed as basic non-continuous local assimilation, that is, not harmony according to our definition. Harmony operating between adjacent syllables (Odden 1994a) is common. In Ndonga (Bantu) (Viljoen 1973) or Lamba (Bantu) (Doke 1938), nasal consonant harmony only occurs when the target and trigger are in adjacent (open) syllables. In vowel harmony, syllable adjacency is difficult to tease apart from blocking and non-iterativity. In Kikuria (Bantu) height harmony, a high vowel causes raising of preceding mid vowels, but an intervening low vowel blocks height harmony. Although Odden (1994a) analyzes this as a case of syllable adjacency, the low vowel could be a blocker, failing to undergo and transmit harmony. Blocking is not a factor in Standard Bengali (Indo-Aryan) harmony (Mahanta 2007). High vowels /i u/ trigger [+ATR] harmony regressively to /ɛ ɔ/ only in the immediately preceding syllable:

(52)	pɔtro	'letter, document'	potrika	'horoscope'
	$k^hɛla$	'game'	k^heli	'to play'
	$kɔt^ha$	'spoken words'	kot^hito	'uttered'
			$kɔt^honiyo$	'speakable'
	pɔd	'position'	pɔdobi	'position holder'

As mentioned in Section 3.1.2, Odden (1994a) proposes that adjacency parameters (root adjacent, syllable adjacent) be added to basic considerations of locality, such that interacting segments have adjacency restrictions imposed, or are unrestricted. Uffmann (2004) recasts Odden's adjacency parameters as optimality-theoretic constraints and Pulleyblank (2002) implements a range of proximal vs. distant

requirements on his sequential anti-disagreement markedness constraints. Kaplan (2007, 2008) has argued that some non-persistent harmony cases are due to minimal satisfaction of positional licensing requirements of harmonic features in specified positions or domains rather than adjacency constraints. For example, in Lango (Nilotic) (Noonan 1992), ATR harmony applies regressively from a suffix to the final root vowel: bɔŋɔ-nɪ → bɔ̀ŋóní 'your dress'. Kaplan treats this as a licensing requirement whereby [+ATR] must be realized on the root; a single affiliation satisfies the constraint.

Two adjacent syllables can constitute a metrical foot, rendering the foot constituent the domain of harmony. In Kera (Pearce 2007a), regressive fronting harmony operates within an iambic foot. Central vowels are fronted only within the same foot as a trigger front vowel: e.g. bàl-é → [(bèlɛ̀ː)] 'love' but bàad-é → [(bàa)dɛ̀] 'wash'. In this case, it is not clear whether stress plays a role in the harmony. However, harmonies do exist that target stressed segments (e.g. metaphony and umlaut patterns), or are triggered by stressed segments, for example, Guaraní nasal vowel-consonant harmony or Old Norwegian height harmony (Majors 1998). In these cases, foot-bounded domains become an issue.

Certain metrical approaches to harmony make reference to asymmetries, for example between heads/non-heads or strong/weak elements (see Halle and Vergnaud 1981 for foundational work). Hualde (1989a) proposes a metrical account of metaphony systems, such as Lena's harmony (Section 2.4.3, Section 3.3.4) (see also Zubizarreta 1979). Hualde's analysis uses the metrical structure constructed for stress assignment: the assimilating feature percolates to targets within it, and the stress foot delimits the margins of harmony. As Majors (1998) points out, not all work in this tradition postulates concidence of metrical stress feet and metrical harmony structure, which loses the advantage of utilizing existing constituents.

Similar issues arise with other foot-based analyses. Flemming (1993) argues that the harmonies in question result from autosegmental spreading rules without reference to stress. Spreading is restricted by a constraint that limits a feature's associations to units within the same metrical foot. Piggott (1996) proposes a similar analysis for Lamba's nasal consonant harmony in adjacent syllables as licensing of the feature by the harmonic foot. This case does not show coincidence with stress patterns, and the "foot" could simply serve as a method of achieving (often) binary groupings of syllables. Likewise, Flemming's approach has been challenged on the basis of the foot structures it requires to obtain the harmony domain (Beckman 1998 [1999]; Majors 1998; Walker 2005).

An area of substantial growth in the last decade centers on licensing approaches to positional asymmetries, which formulate constraints in terms of position-sensitive faithfulness or markedness. Beckman (1998 [1999]) proposes stressed-syllable faithfulness constraints for nasal harmony in Guaraní, which is triggered by stressed nasal vowels and blocked by stressed syllables that contain an oral vowel. Beckman uses a faithfulness constraint for [nasal] in stressed syllables, Ident-σ(nasal), which, together with a markedness constraint that drives harmony, captures both the triggering and blocking status of stressed syllables.

A positional markedness approach to stress sensitive harmonies is developed by Majors (1998) and Walker (2005), which emphasizes hypothesized articulatory and perceptual bases of these patterns. Their analyses employ a licensing constraint that requires a given feature specification to have an association to a stressed syllable (cf. Klein 1995). The constraint is satisfied either when the feature specification is only linked to a stressed position or is linked to both stressed and unstressed syllables, a configuration known as *indirect licensing* (Steriade 1995). Majors teams this constraint with a faithfulness constraint for stressed syllables to obtain harmony patterns where segments in unstressed syllables assimilate to stressed ones. In harmonies where a stressed vowel assimilates to an unstressed one, stressed-syllable faithfulness is dominated by another constraint that determines control by an unstressed trigger. Examples include a morpheme-specific faithfulness constraint for harmonies triggered by particular inflectional vowels (Majors 1998) or a phonological constraint that blocks formation of the vowel quality that would occur under assimilation of the unstressed vowel to the stressed vowel (Walker 2005).

4.2.2 Morphological/Morphophonological Domains A standard domain of harmony is the word, in which harmony applies across internal morpheme boundaries. The "word" may correspond to the morphological notion of word, or be described as the "phonological word," a prosodic constituent, if clitics are included.[33] In fact, vowel harmony is often used as a diagnostic for determining word boundaries (Suomi, McQueen and Cutler 1997; Bauer 2003: 60). There are nevertheless cases in which harmony is restricted to the root or behaves differently within the root, and others in which certain affixes are non-undergoers of harmony.

Various consonant harmony patterns are confined to the root, including laryngeal, nasal, and dental harmony. Ngbaka (Adamawa-Ubangi) is cited as a root-restricted vowel harmony pattern (Archangeli and Pulleyblank 2007). In some languages, roots show disharmonic patterns while harmony applies across morpheme boundaries, as argued by Clements and Sezer (1982) for Turkish. Pulleyblank (2002) also makes a case for a root/word domain distinction based on differing patterns in the height harmony system of C'Lela (Benue-Congo) (Dettweiler 2000), in which the sequence high vowel non-high vowel is unattested in roots, but is possible spanning the root-suffix boundary.

Further cases exist where affixes or clitics may fail to harmonize. This may be due to idiosyncratic reasons or their peripheral status in the word. In Wolof (Atlantic) (Ka 1994), progressive ATR harmony changes the vowel /a/ to [ə]: nɛlaw-a̱m 'his/her sleep' versus dugub-ə̱m 'his/her millet'. The agentive suffix /-kat/ fails to become [+ATR] when associated to [+ATR] stems: luxus-kat 'magician' *[luxus-kət]. It can also initiate a new [−ATR] harmonic domain: luxus-kat-am 'his/her magician'. The standard approach to these cases is to specify the segment/morpheme with an underlying [−ATR] specification.

In Standard Yoruba, subject clitics do not harmonize with the root. In Oyo and Ibadan Yoruba, back round subject clitics do harmonize (Akinlabi and Liberman 2000). Przezdziecki (2005) treats clitics as part of the phonological word.

(53)	Standard		Oyo/Ibadan		
	mo	lɔ	mɔ	lɔ	'I went'
	mo	yó	mo	yó	'I am fed'
	ó	wá	ɔ́	wá	'he/she came'
	ó	dé	ó	dé	'he/she came'

The domain of vowel harmony does not always match the prosodic domains of other phonological processes, however, and may be difficult to define morphologically, as some clitics participate and some do not. Kabak and Vogel (2001) conclude that domains defined in terms of prosodic constituents such as the "phonological word" or "clitic group" do not accurately denote the domain of vowel harmony in Turkish.

In Bantu, vowel harmony typically operates within a morphological domain consisting of the verb stem minus the final vowel (Hyman 1999) and does not extend to pre-verb stem clitics. The verb stem in Bantu does not correlate exactly with a derivational/inflectional split, since it contains some inflectional morphemes. Terms such as the *macrostem* or *extended stem* have been proposed (e.g. Myers 1987 [1990]).

One method of referencing domains in OT is through morphologically indexed constraints. This is achieved by having versions of the same constraint subscripted for domains. For Kinande vowel harmony, Archangeli and Pulleyblank (2002) propose domain specific versions of Hɪ/ATR: $\text{Hɪ/ATR}_{\text{Root}}$ >> $\text{Hɪ/ATR}_{\text{stem}}$ >> $\text{Hi/ATR}_{\text{macrostem}}$ to achieve the blocking effects in Kinande described in Section 4.1. Other constraints, such as Align-L (responsible for regressive harmony) are ranked above and between them. Archangeli and Pulleyblank (2002) further argue that constraints differentiated by domain are harmonically ranked from smaller to larger domain, essentially achieving the cyclic effect of harmony operating from the root outwards.

Harmony can apply to domains larger than the word. Ka (1994) argues that vowel harmony in Wolof applies within the phonological phrase. This includes the head of the phrase (noun or verb) plus complements to the right, ex. [dugg ngə cə] 'you went into it' versus [dɛm nga ca] 'you went to it' or [[goor ñə dinəñu ko] gas] 'the men will dig it' versus [[xalɛ ya dinañu kɔ] door] 'the children will hit him/her'. Phonological phrases do not always correspond to syntactic structure. Harmony domains that cross word boundaries are also reported for Nawuri (Casali 2002), Somali (Cushitic) (Hall *et al.* 1974), and Vata (Kru) (Kaye 1982).

Finally, there may be optionality in whether harmony applies, and gradience in the extent of the harmony. Mutaka (1995) observes that harmony in Kinande in a phrase such as èmítí míkù:hì 'short trees' can affect no preceding vowels, one [èmítí míkù:hì], two [ɛ̀mítí míkù:hì], or all vowels [èmítí míkù:hì]. The further away from the trigger a morpheme is, the less likely it is to harmonize. This can be viewed in more functional terms if harmony is analyzed as extension of gestures, and the 'strength' of the gesture fades the further it is from the original source. This would suggest that vowels are less strongly altered further away, apparently the case for Kinande (Archangeli and Pulleyblank 2002).

In summary, the role of morphological domains is generally recognized and incorporated into analyses, but there has been little debate on how the extent of harmony in terms of morphology should be addressed. In OT, the issue is addressed by indexing constraints for the domains in which they operate (Archangeli and Pulleyblank 2002).

5 Shifts in Empirical Focus

Linguistic analysis rests on an empirical foundation. Recent empirical work has shown a shift in the types of data being emphasized. Four particular categories are (i) research on lesser-studied languages, (ii) instrumental studies, (iii) studies of variation and/or statistical tendencies and (iv) psycholinguistic production tasks. Research in these directions has led to refinements in our understanding of harmonic issues and have brought about new theoretical advances.

The importance of research on lesser-studied languages is reflected in the variety of languages discussed here, and in the contribution of new data to typological generalizations and theoretical claims. As discussed in Section 4.1, certain research has pushed to make directionality effects in harmony derivative of other properties of the system. However, descriptive work of under-studied languages, such as Karajá (Ribeiro 2003) and Tuvan (Harrison 2000), has provided solid evidence to the contrary. Lacking these studies, the question of whether directionality exists as an independent characteristic of vowel harmony would be more ambiguous. Furthermore, descriptions of endangered languages can reveal illuminating changes in harmony systems. Anderson and Harrison (to appear) present a study of Tofa, a moribund Turkic language, in which vowel mergers have taken place, creating a more abstract vowel harmony system for younger speakers, as well as considerable micro-variation in round vowel harmony. Therefore, it is essential that detailed description of harmonies in languages be pursued, going beyond the well-known systems that have formed the primary emphasis of research to date.

Experimental studies of variation or statistical tendencies have been conducted for the vowel harmonies of Hungarian and Finnish. While both languages have backness harmony with harmonic and neutral vowels, closer examination of speakers' behaviors revealed subtleties not accurately captured in previous descriptions. Ringen and Kontra (1989) performed a questionnaire-based study on Hungarian that investigated suffix vowel choice with disharmonic roots (mostly loans) ending in *neutral* front vowels [i], [iː], [e], and [ɛ], which are reported to be transparent to backness harmony. They discovered that the lowest neutral vowel [ɛ] actually triggers harmony in most cases, the second lowest vowel [e] tends to be transparent but shows some variability, and the highest neutral vowels are indeed transparent. The study identified considerable vacillation in suffix vowel choice following sequences of two syllables with neutral suffix vowels. A connected study by Kontra, Ringen, and Stemberger (1991) found that sentence context influences suffix vowel choice in words that show vacillation. In more

recent research, suffix vowel choice with stems containing neutral vowels has been investigated by Hayes and Londe (2006), using a 'wug' test, where speakers selected suffix forms for novel stems. They also collected data on quantitative patterns using a web search approach (see also Wanlass 2008 for an online corpus study).

A study of Finnish by Ringen and Heinämäki (1999) also investigated suffix vowel choice with disharmonic loanwords using questionnnaires. In loans, harmonic front and back vowels were reported to pattern asymmetrically, with certain normally harmonic front vowels behaving as transparent. Ringen and Heinämäki's study found that in disharmonic roots where the last vowel was back, the suffix vowel was almost always back, that is, the final vowel was a trigger. In disharmonic roots where the final vowel was front, many forms presented variation, determined by a variety of factors such as stress and vowel quality. Not only have these studies uncovered aspects of the harmony systems that were hitherto unknown, but also the statistical tendencies that they identify are problematic for standard rule-based or classic OT approaches, necessitating revisions to the theory (Ringen and Heinämäki 1999; Hayes and Londe 2006).

New research on harmony has also emerged in the field of artificial language learning and experimental production tasks, research which tests the naturalness and functional underpinnings of harmony systems. Pycha *et al*. (2003) trained naïve speakers on different patterns of non-local vowel interaction, both harmony and disharmony, and Wilson (2003) tested adults' ability to learn nasal consonant harmony or disharmony patterns in suffix choice. Both studies concluded that speakers learned the harmonic/disharmonic systems, but did not learn more "random" or complex rules. Mintz and Walker (2006) tested English-learning infants' sensitivity to vowel color harmony using the head-turn preference procedure. The infants showed an ability to segment words based on color harmony even though their ambient language environment had not previously exposed them to vowel harmony patterns. Koo and Cole (2006) tested liquid consonant harmony/disharmony versus back vowel harmony/disharmony and found that liquid (dis)harmonies were more easily learned. They concluded that this was due to the perceptual similarity involved in liquids, as highlighted in recent work on consonant harmony. Other experimental learning-based studies of harmony are reported by Finley (2008) and Zaba (2008). Research by Cole *et al.* (2002) tested speech production (production time and error rate) for nonce words in which vowels agreed on the front/back dimension versus the height dimension. They found that front/back harmony facilitated speech production but height harmony did not. Walker (2007) and Rose and King (2007) used different speech error elicitation tasks to test connections between similarity and speech production underlying the analysis of consonant harmony systems. Rose and King (2007) examined the impact of harmony constraints on speech errors, and found elevated speech error rates for sequences that violated laryngeal harmony. Walker (2007) found that the consonants that were more prone to participate in speech errors with nasals in English matched the ones preferentially targeted in nasal consonant harmony across languages.

Instrumental studies of harmony have been a critical source of new evidence on topics that have long been the subject of debate. They have been especially valuable on issues that cannot be straightforwardly resolved by ear-based transcription. Acoustic studies of faucal harmony and emphasis harmony have been conducted by Bessell (1998) and Shahin (2002), shedding new light on the properties of these tongue root systems. Articulatory studies have also proved especially revealing. A study of Kinyarwanda by Walker *et al.* (2008) used electromagnetic articulography to uncover evidence that the harmonizing retroflex posture persists during reportedly transparent non-coronal consonants when they occur between audibly harmonizing fricatives. A lingual ultrasound study of Kinande by Gick *et al.* (2006) found that /a/, reported to be transparent in the language's harmony, actually shows advanced and retracted root positions consistent with its full participation in tongue root harmony. Research on Hungarian by Benus and Gafos (2007) using electromagnetic midsagittal articulometry examined neutral front vowels. They found that the neutral vowels in monosyllabic stems that select front vowel suffixes had a significantly more advanced tongue position than ones in stems that select back vowel suffixes.[34] However, the difference in tongue advancement in these vowels did not alter their front perceptual quality, accounting for its failure to be reflected in transcription. The question of when a subtle but consistent degree of shift in articulation reaches the criterion for a difference in the phonological representation of segments is not uncontroversial. Thus, instrumental research can clearly contribute on various outstanding issues in harmony on a case-by-case basis. At the same time, it raises new questions for the goals of phonological analysis and how and whether observations of data involving variation along a phonetic continuum should be modeled within phonology.

Across each of the above categories of data types, much remains to be documented and discovered. Continued study in these directions will surely continue to shed new light on the topic of harmony.

6 Conclusion

This chapter has attempted to synthesize and elucidate the main contributions of recent research into harmony systems. There has been considerable progress made over the last decade in the study of harmony, but at the same time divergence of analysis. Of course, it is hard to do justice to such a vast topic in an overview chapter. Within the bounds that this paper affords, we have concentrated on certain themes and theoretical approaches, but there remain areas of research that are worthy of consideration beyond that which we can cover here. Nevertheless, we are confident that the topic of harmony is sufficiently rich that readers will be able to use this chapter as a platform to explore topics in greater detail and make their own future contributions to the study of harmony systems.

NOTES

1 Hansson (2001b, 2010) defines consonant harmony as operating over at least a vowel. Indeed, consonant harmonies rarely apply across a string of two or more consonants, but such cases do exist. In Imdlawn Tashlhiyt Berber (Elmedlaoui 1992), sibilant harmony can apply across strings that lack a vowel: e.g. s-bbrbʃ → [ʃbrbʃ] 'be gaudily colored' (underlining indicates pharyngealization).
2 We do not discuss consonant harmony in child language, which does involve harmony for place. See Goad (1997), Rose (2000), Pater and Werle (2001, 2003) for recent discussion.
3 Although Harms (1985) transcribes this form with a pre-nasalized [s], Harms (1994) has more recently indicated that [s] is not pre-nasalized after a nasal vowel.
4 Harms (1985) states that /s/ blocks progressive nasal spreading. The more recent description by Harms (1994: 8) seems to suggest that /s/ does not invariably block spreading, but among the data provided in that work he includes the example [mĩə̃su] 'spear', in which it serves as a blocker.
5 Shahin (2002) argues that St'át'imcets Salish (Lilloet) has a post-velar harmony like Arabic, but as it affects only a single adjacent segment to the left, this does not fall under our definition of harmony and would be analyzed as local assimilation.
6 Walker *et al.* (2008) also find evidence that points to [ɽ] undergoing harmony.
7 Tuvan also shows a round harmony that we do not discuss here.
8 Whether pharyngeal harmony is formally distinct from RTR harmony is an open question. For discussion, see Li (1996: 53), Svantesson *et al.* (2005: 8); note also Casali (2003).
9 In inflectional suffixes, /u/ is the only contrastive high vowel and the only trigger for height harmony. See Dyck (1995) and Campos-Astorkiza (2007) for discussion.
10 Paradis (1992) posits only five phonemic vowels.
11 Features may also belong to the root node itself, as has been suggested for major class features (Schein and Steriade 1986; McCarthy 1988).
12 A NoGap constraint (or its equivalent) has been posited within Con, that is, the set of rankable constraints that compose an optimality-theoretic grammar. In some analyses it is undominated (Itô, Mester and Padgett 1995; Padgett 1995a), whereas in others it is violable within the patterns under study (e.g. Smolensky 1993; Uffman 2004).
13 Some of the other harmonies Shaw proposes, such as labial, are dissimilatory morpheme structure constraints or morphological affixation, rather than true consonant harmony.
14 This reasoning would not work for feature systems in which coronals and front vowels share specification (Clements and Hume 1995).
15 Hansson (2007b) has argued that the speech production explanation is not valid for all consonant harmonies, particulary those with secondary articulation. Those cases have unique diachronic explanations, often due to language contact and related to (re)interpretation of C-V coarticulation.
16 The model of aggressive reduplication (Zuraw 2002) employs a similar mechanism to couple substrings, but does not encode similarity directly.
17 In Rose and Walker (2004), directionality is added to the Ident-CC constraint.
18 Hansson (2007a) argues that while lack of blocking is a descriptive characteristic of consonant harmony systems, it does not necessarily follow from the agreement-by-correspondence approach. See Hansson (2007a) for discussion of two scenarios under which blocking might arise.

19 Krämer (2001, 2003) develops a surface correspondence approach for vowel harmony, but builds adjacency at a moraic or syllabic level into the definition. Pulleyblank (2002) offers a different perspective that accounts for both vowel and consonant harmony using a "no-disagreement" harmony-driver (see also Archangeli and Pulleyblank 2007).
20 Cole and Kisseberth (1995a, b) propose an alignment-driven analysis of vowel harmony and nasal vowel-consonant harmony that posits *feature domains* in place of traditional autosegmental representations.
21 Other proposals have been made to treat some or all patterns of vowel harmony as driven by identity constraints among vowels which stand in a correspondence relation in the output (Kitto and de Lacy 1999; Krämer 2003).
22 Retroflex harmony in Kinyarwanda could prove a challenge for such analyses. Walker *et al.* (2008) find articulatory evidence that is consistent with an interpretation of the pattern as a regressive vowel-consonant harmony. However, the harmony appears to occur only when it extends to particular target consonants. Non-coronal consonants that precede a retroflex fricative trigger only show evidence of undergoing harmony in the presence of a preceding target, a fricative or flap.
23 Boersma (1998) also notes that reduction of articulatory contours could produce vowel harmony patterns.
24 See Baković (2000) for discussion of further issues presented by feature-driven markedness.
25 See Gick *et al.* (2006) for related acoustic research on the harmony system of Kinande (Bantu).
26 For other proposals with some related concepts, see Schane (1990) and Casali (1996 [1998]).
27 See Clements and Osu (2003) for an alternate perspective on defining the hierarchy.
28 Related issues for the treatment of glides arise surrounding their capacity to block nasal vowel-consonant harmony in some languages but not others. For discussion, see Cohn (1990, 1993b), Walker (1998 [2000]), Flemming (2004), Levi (2004, 2008), and Padgett (2008).
29 Lango vowel harmony is subject to a number of other complex conditions, including the condition that two intervening consonants do not block progressive harmony if the trigger vowel is high.
30 See Mester (1988) for a solution to identity effects in terms of dependent tier ordering.
31 An acoustic study by Gordon (1999a) of backness harmony in Finnish leads him to conclude that harmony functions at a low phonetic level for "transparent" /i/ and /e/. Svantesson *et al.* (2005) characterize /i/ as phonologically transparent to pharyngeal harmony in Halh, but their acoustic examination of this vowel reveals that it is realized as pharyngealized in pharyngeal words.
32 Such a system does occur with vowel-consonant emphasis harmony, but apparently not with vowel harmony.
33 Compound words may be considered a morphological word, but typically consist of two distinct harmonic domains, as in Finnish, Hungarian, or Turkana. Hoberman (1988) reports that in Azerbaijani Jewish Aramaic, emphasis harmony may sometimes extend to the other half of a compound. Suffixes added to compounds harmonize with the second half, forming a phonological word domain with the second portion which does not coincide with the morphological relationship of the suffix attaching to the whole compound.
34 The vowels were measured in monosyllabic stems without a suffix, which prevents vowel-to-vowel coarticulation from affecting the result.

9 Contrast Reduction

ALAN C. L. YU

1 Introduction

The notion of *contrast reduction*, which encompasses both notions of *mergers* and *neutralization*, presupposes the concept of contrast. Two sounds are phonologically contrastive if they are in opposition with each other, that is, if they are capable of signaling a meaning difference between two lexical items in a particular language (Steriade 2007; Dresher 2009). The plosives [p] and [p^h], for example, are in opposition with each other in Cantonese (Yue Chinese) (e.g. [paː˥] 'father' vs. [p^haː˥] 'to lay face-side down') but [b̥] and [p] are not. Contrast is not restricted to pairs of segments; classes of segments contrast as well. The aspiration opposition between [p] and [p^h] finds analogs in other pairs of segments (t~t^h, k~k^h, k^w~k^{wh}). When a phonological opposition is suspended, neutralization or merger obtains. For example, Cantonese has no aspiration opposition between plain and aspirated plosives in syllable-final position; all syllable-final plosives are voiceless and unreleased (e.g. [t^haːp˥˧] 'pagoda', [paːt˥˧] 'eight', [kɔːk˥˧] 'corner, horn').

The terms *merger* and *neutralization* are often employed in complementary contexts; *merger* often characterizes a diachronic collapse of contrast, neutralization a synchronic. The diachronic-synchronic divide between merger and neutralization is more apparent than real, however; the two notions are the two faces of the same coin. The notion of merger is often applied in the context where a contrast reduction leaves no trace of the said contrast in the synchronic system, of which a context-free contrast reduction is a clearest example. Neutralization applies to contrast reduction that is context-dependent; traces of a contrast remain in some contexts, but not in others. Certain varieties of English, for example, merges the voiceless labio-velar fricative /ʍ/ with its voiced counterpart /w/ (Minkova

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

2004). Thus the words *whine* and *wine* are homophonous; no remnant of this /ʍ/~/w/ contrast is evidenced in the grammar of speakers of these dialects. In certain dialects of Cantonese (most prevalently in Guangzhou, Hong Kong, and Macau; Bauer and Benedict 1997), the distinction between plain versus labial velars is not maintained before the back rounded vowel /ɔ/. The collapse of the plain vs. labial velars distinction is referred to as a matter of neutralization because the contrast remains before vowels that are not /ɔ/ (e.g. [kɐn˥] 'tight' vs. [kʷɐn˥] 'boil'). These instances of contrast reduction in English and Cantonese transpire diachronically, but one results in a merger (i.e. the /ʍ/~/w/ merger) and the other in neutralization (i.e. k(ʰ)~kʷ(ʰ) neutralization). In this chapter, I shall assume that the term *neutralization* refers to contrast reduction that results in alternation, while the term *merger* will refer to any reduction of contrast, both synchronically and diachronically. Thus, in the case of the kʰ~kʷʰ contrast in Cantonese, kʰ~kʷʰ *merge* before /ɔ/ diachronically. The outcome of this *merger* is the *neutralization* of /kʰ/ and /kʷʰ/ before /ɔ/.

This article begins with a review the range of contrast reduction (Section 2). Section 3 surveys several theories that attempt to explain the sources of contrast reduction. Section 4 concludes with a discussion of the challenges to a purely phonological conception of contrast reduction.

2 Typology of Contrast Reduction

Contrast reduction manifests itself in three different ways: structure-preserving reduction, structure-building reduction, and free variation.[1] *Structure-preserving reduction* characterizes scenarios where two or more distinct sounds have, after the reduction, a constant form that is physically similar to that of one of the sounds appearing in the position of differentiation (e.g. the kʰ~kʷʰ neutralization; cf. Kiparsky 1985). Formally, a reduction of contrast *m* is structure-preserving if and only if *m* turns two (or more) distinct sounds into only one of the two sounds, to the exclusion of the other. The merger of /ʍ/ and /w/ is structure-preserving since the result of the merger leaves /w/ as the surviving sound. Regressive assimilation of voicing is another instance of structure-preserving contrast reduction. For example, in Dutch, the distinction between voiced and voiceless plosives is suspended preconsonantally (Ernestus and Baayen 2003). However, the result of neutralization differs depending on the nature of the following consonant. For example, before a voiced plosive, the /t~d/ contrast in *verwijten* [vɛrʊɛitən] 'reproach-INF' and *verwijden* [vɛrʊɛidən] 'widen-INF' neutralizes toward /d/ (*verwijt bijna* [vɛrʊɛid bɛinaː] 'reproach almost' vs. *verwijd bijna* [vɛrʊɛid bɛinaː] 'widen almost'). However, before a nasal, neutralization is toward /t/ (*verwijt niet* [vɛrʊɛit nit] 'reproach not' vs. *verwijd niet* [vɛrʊɛit nit] 'widen not').

Contrast reduction is structure-building when the outcome of contrast reduction is a sound intermediate between the normal realization of the two phonemes. The case of final-consonant voicing neutralization in Cantonese is a case in point. Stops in syllable-final position are unreleased, thus phonetically do not contrast

in terms of aspiration. Another celebrated case of structure-building reduction is flapping in English. The /t~d/ contrast in English is suspended intervocally where the coronal in question is immediately followed by an unstressed vowel (e.g. *heed* ['hiːd] vs. *heat* ['hiːt], but *ladder* ['læɾɚ] vs. *latter* ['læɾɚ].[2]

When contrast reduction leads to a form varying between two or more variants, this is referred to as *free variation.* For a large number of Cantonese speakers, syllable-initial /n/ is in free variation with /l/ (Bauer and Benedict 1997). Thus, words like [nejˊ] 'you' and [naːnˋ] 'difficult' are often pronounced with initial /l/, thus merging with [lejˊ] 'Li (surname)' and [laːnˋ] 'orchid' respectively. The rate of /n/ vs. /l/ usage varies according to age and gender of the speaker as well as the register of speaking (e.g. read speech vs. conversational speech).

2.1 Positions of Contrast Reduction

Contrast is often restricted to certain positions within the word: the syllable peak (rather than the margin), the onset (rather than the coda), the stem (rather than the affix), the stressed syllable, or the edge of the morphological domains. Washo (Hokan), for example, only allows voiceless liquids and nasals in onset position (Jacobsen 1964). Isthmus Zapotec (Oto-manguean) contrasts glottalized and modal-voiced vowels, but only in stressed positions (Bueno-Holle 2009). Hausa (Chadic) has a five-vowel system (i, e, a, o, u) with a long-short distinction which is reliably distinguished only in final position (Steriade 1994). Ngalakan (Australian) has a five vowel system (i, e, a, o, u). Mid vowels in Ngalakan are restricted to the edges of roots (Baker 1999: 72–73); if there is only one mid vowel in a root, it must appear in an edgemost syllable (i.e. initial /c**e**raṭa/ 'woman's ceremony' or final /curuw**e**-/ 'rush'). If there is more than one mid vowel, they must occur in contiguous syllables (/caw**oro**/, 'patrilineal clan') or every vowel in the root must be a mid vowel (/k**o**w**ele**ŋʔ(–mi+)/ 'beckon to'). !Xóõ (Bushman), contrasts consonants with clicks and consonants without click accompaniment, but only in initial syllables (Traill 1985). In Etung (Bantu), falling and rising tones (H͡L, H͡H], L͡H) are restricted to the final syllable of phonological words, but there is no restriction on the occurrence of level tones (Edmondson and Bendor-Samuel 1966). In Lushootseed (Central Salishan), glottalized consonants are only found in roots and lexical suffixes; grammatical suffixes never have glottalized consonants (Urbanczyk 1996: 46). Contrast restrictions might also differ across word types. For example, in a cross-linguistic survey of 32 languages having twenty-six consonants or more, Willerman (1994) found that pronouns made significantly less use of the palato-alveolar, retroflex, uvular, and pharyngeal places than other places of articulation, and fewer laterals, affricates, trills, clicks, ejectives, and aspirated segments.

Loci of contrast reduction are not always characterizable in structural terms. Steriade (1994) observes that languages with a retroflexion contrast in the apicals (e.g. t vs. ṭ) often neutralize the contrast in initial or post-consonantal positions, but allow the contrast in post-vocalic position. The position of retroflexion neutralization is difficult to capture in prosodic terms since post-vocalic position can

be either within or across a prosodic domain (e.g. the coda of a syllable and a syllable onset in intervocalic position). Obstruents in Lithuanian contrast in terms of voicing (Senn 1966; Steriade 1999b). However, the voicing contrast is supported only before sonorants (*sko**bn**is* 'table'; *bã**dm**etys* 'year of famine') and not elsewhere. Voicing is neutralized word-finally (*daũ**g*** [dau**k**] 'much'; *kà**d*** [ka**t**] 'that') and in pre-obstruent position (*dè**g-t**i* [**kt**] 'burn-inf', *míela**s** **d**raũgas* [zd] 'dear friend').

2.2 *Common Triggers and Targets of Contrast Reduction*

Languages with contrast reduction often exhibit striking parallelism in the direction of merger and neutralization. Non-assimilatory neutralization of laryngeal contrasts in word-final and pre-consonantal positions is often structure-preserving; the preserved segments are generally voiceless. Neutralization toward voiced or ejective is rare if not non-existent.[3] Reduction of vocalic contrasts in unstressed positions is commonplace across the world's languages. The vast majority of such reductions involves the neutralization of vowel nasalization, quantity, or height. Nasal and oral vowels, for example, are often only contrastive in the stressed syllables (e.g. Copala Trique (Hollenbach 1977), Guaraní (Beckman 1998: 158)). Contrasts in vocalic quantity are frequently neutralized toward the short variant in unstressed syllables. Kolami, for example, only contrasts long and short vowels in initial syllables, which are always stressed (Emeneau 1961: 6–7). Quantity contrasts may also neutralize toward the long variant under certain circumstances. For example, a vowel following a consonant-glide sequence must be long (***ak**-a* 'ask!' vs. *kw-**a:k**-a* 'to ask'; Myers and Hansen 2005: 318) in Kinyarwanda (Bantu), which has a contrast in vowel length ([ɡusiːβa] 'to be absent' vs. [ɡusiβa] 'to erase'; Kimenyi 1979: 1). Reduction in vowel height in unstressed position often favors one of two outcomes: the unstressed vowel may become either [a] or [ə]. In Belarusian, for example, mid vowels /e, o/ reduce to [a] (['n**o**ɣi] 'legs' vs. [n**a**'ɣa] 'leg'; ['r**e**ki] 'rivers' vs. [r**a**'ka] 'river'; Crosswhite 2004: 192); thus the five vowels found in stressed syllables, /i, e, a, o, u/, are reduced to three, [i, a, u], in the unstressed syllables. The seven-vowel system in Central Eastern Catalan (/i, e, ɛ, a, ɔ, o, u/) is only evident in stressed syllables; in unstressed syllables, only three vowel qualities ([i, ə, u]), are allowed; underlying /e, ɛ, a/ become [ə] while /u, o, ɔ/ become [u] (1). Vocalic contrast reductions along other featural dimensions are rare and are often secondary to height neutralization in the same system (Barnes 2002).

(1) Central Eastern Catalan (Barnes 2002: 37)

řiw	'river'	řiwɛ́t	'river, dim.'
néw	'snow'	nəwɛ́tə	'snow, dim.'
mɛ́l	'honey'	məlɛ́tə	'honey dim.'
pálə	'shovel'	pəlɛ́tə	'shovel, dim.'
řɔ́ðə	'wheel'	řuðétə	'wheel, dim.'
mónə	'monkey fem.'	munɛ́tə	'monkey, fem. dim.'
kúrə	'cure'	kurɛ́tə	'cure, dim.'

The targets of assimilatory neutralization show cross-linguistic similarities as well (Cho 1990; Ohala 1990; Jun 1995; Steriade 2001; de Lacy 2002, 2006). For example, obstruents are often voiced after nasals (Pater 1999). Nasals in turn frequently assimilate to the place of articulation of the following consonant, as illustrated by the examples from Yoruba (Niger-Congo) in (2).

(2) Yoruba nasal assimilation (Pulleyblank 1995: 5)

a.	bá	m̀bá	'overtake'
	fɔ́	ŋ̀fɔ́	'break'
b.	tà	ǹtà	'sell'
	sũ	ǹsũ	'sleep'
c.	jó	ɲ́jó	'dance'
	jɛ	ɲ́jɛ	'eat'
d.	kɔ	ŋ́kɔ	'write'
	wí	ŋ̀wí	'say'
e.	gbɔ́	ŋ̀mgbɔ́	'hear, understand'
	kpa	ŋ̀mkpa	'kill'

Among obstruents, coronals are most susceptible to place assimilation. In Korean, for example, morpheme-final coronals assimilate to dorsals or labials (3a). Morpheme-final labials assimilate to dorsals (3b), but no assimilation is observed when the following consonant is coronal. Dorsals are most inert; they assimilate neither to a following labial nor to a following coronal (3c).[4]

(3) Korean place assimilation (Hume 2003: 7–8)[5]

a.	/mit+ko/	[mikk'o]	'believe and'
	/mit^{h}+pota/	[mipp'ota]	'more than the bottom'
b.	/ip+ko/	[ikk'o]	'wear and'
	/nop+ta/	[nopt'a] *[nott'a]	'high'
c.	/nok+ta/	[nokt'a] *[nott'a]	'melt'
	/kuk+pota/	[kukp'ota] *[kupp'ota]	'more than soup'

3 Theories of Contrast Reduction

Early discussions of contrast reduction focused on how to characterize the outcome of context-specific contrast reduction. That is, how would a theory of phonemics capture the fact that the contrast between two or more sounds in some positions of a word or a syllable is not maintained in some other positions? The main analytic puzzle neutralization presents to structuralist phonemics concerns the violation of the bi-uniqueness condition (i.e. the non-one-to-one mapping between allophones and phonemes). The Prague School resolves this indeterminacy by positing archiphonemes in contexts of neutralization (Trubetzkoy 1939); archiphonemes are units that represent the common features of phonemes whose

contrastive property is neutralized in specific contexts. In Yoruba, for example, a preconsonantal nasal would be treated as an archiphoneme, N (e.g. ḿbá /Ńbá/ 'overtake'). The archiphonemic treatment of neutralization anticipates the underspecification treatment of neutralized segments made possible by the reconceptualization of the phonemes as sets of distinctive features. Under an underspecification model, a preconsonantal nasal in Yoruba, for example, would be specified for the feature [+nasal], while the surface realization of this underspecified nasal would be specified contextually.

In addition to the issue of representation, theories of neutralization also attempt to explain the causes for neutralization. That is, why do cross-linguistic parallelisms abound in cases of contrast reduction? Two main approaches have been advanced: structure-based and cue-based. This section reviews how these two approaches conceptualize the problem of contrast reduction and what mechanisms account for the observed typological tendencies.

3.1 Licensing and Markedness

Structure-based approaches maintain that certain prosodic or structural positions disfavor the maintenance of phonological contrasts. The phonological grammar may either prohibit a contrast in a given structural position in terms of a filter constraint (4) or impose a licensing condition which specifies how a phonological contrast must be configured in order to be realized in a given position within the word (5).

(4) Positional neutralization: filter/negative version (Steriade 1995: 120)
**αF* in *x* where *x* is defined prosodically or morphologically.

(5) Positional neutralization: licensing/positive version (Steriade 1995: 121)[6]
αF must be licenced in *x*, where *x* is defined prosodically or morphologically.

Codas in Pali, for example, must be the first half of a geminate structure (6a) or nasals (6b). Coda nasals must be placeless or homorganic with the following stop.

(6) Pali cluster simplification (Zec 1995: 157)

a.	sup+ta	sutta	'to sleep'
	tap+ta	tatta	'to shine'
	caj+ta	catta	'give out'
b.	dam+ta	danta	'to tame'
	vam+ta	vanta	'to investigate'

Coda constraints such as those in (7) prevent the admission of illicit codas. (7a) states that "if there is a syllable-final consonant which is singly linked, its melody cannot be [–nas]"; (7b) states that "if there is a syllable-final consonant which is singly linked, its melody must be [+nas]."

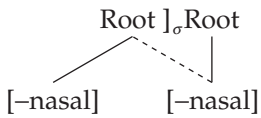
(7) Codas in Pali (following Itô 1986)

a. *C]$_\sigma$ b. C]$_\sigma$

[–nasal] [+nasal]

Geminates, where the melody is doubly-linked both to the coda of one syllable and to the onset of the following syllable, violate neither (7a) nor (7b) because the melody is not uniquely linked to a [nasal] feature.

(8)

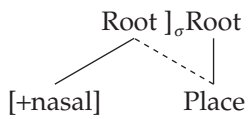


The same approach can be applied to the fact that codas in Pali are either placeless, as in the case of nasal codas, or homorganic with the following stop (9).

(9) Coda place in Pali (following Itô 1989: 224)

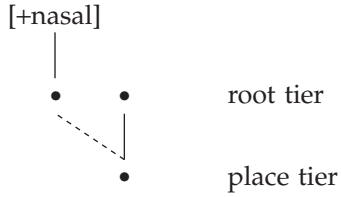
A coda consonant can be specified for place as long as the Place node is not uniquely linked to the coda consonant. If a coda nasal cannot share Place with another segment, it will remain placeless.

(10)



The restrictiveness of potential triggers and targets of neutralization have provided fruitful venues for discovering the organization of features at the phonological level. Many have proposed organizing features into a hierarchical set structure within Autosegmental Phonology (see McCarthy 1988; Clements and Hume 1995 for overviews of proposals in feature geometry). By assuming the different features for place of articulation are hierarchically linked to a Place node, nasal place assimilation in Yoruba can be elegantly and economically modeled in terms of the spreading of the place node (11).

(11) Place assimilation in a feature-geometric organization (Pulleyblank 1995: 9)



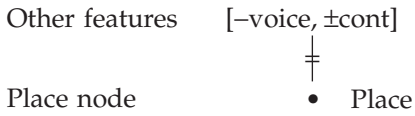
Within this type of feature-geometric framework, non-assimilatory contrast reductions are generally treated as a matter of the delinking of branches of a feature tree. In the Kelantan dialect of Malay (Austronesian), for example, /p, t, k/ neutralize to [ʔ] and /s, f/ become [h] (12).

(12) Kelantan Malay place neutralization (Teoh 1988)

/ikat/	ikaʔ 'tie'	/səsak/	səsaʔ 'crowded'
/dakap/	dakaʔ 'embrace'	/hampas/	hapah 'husk'

Debuccalization to [ʔ] and [h] can be viewed as delinking of the *place* node (13). The fact that /p, t, k/ debuccalize to [ʔ], but /s/ to [h] can be attributed to the fact that non-place features of the underlying segment (e.g. [cont]) are left intact.

(13) Formalization of $s \rightarrow h$ and $p, t, k \rightarrow ʔ$



Adopting the framework of Optimality Theory (OT; Prince and Smolensky 2004), which determines the contrastive status of a feature F via the interaction of a constraint that requires the preservation of F and constraints on the rest of the system (Kirchner 1997), Lombardi (2001) analyzes place neutralization, such as (12), in terms of the interaction between consonantal place faithfulness and a family of universally ranked place markedness constraints ((14); cf. Prince and Smolensky 1993; Smolensky 1993). Unlike the position-specific markedness constraints in (4) and (5), this family of markedness constraints captures the idea that pharyngeals, including /ʔ, h/ (McCarthy 1994), are less marked than coronals in general, irrespective of position. The tableau in (15) illustrates a markedness-based treatment of coda place neutralization.

(14) Place hierarchy: *Dor/*Lab >> *Cor >> *Phar (Lombardi 2001: 29)

(15) Place neutralization in Kelantan Malay (Lombardi 2001: 31)[7]

/ikat/	Max	Dep	*Dor/*Lab	*Cor	*Phar	*MaxPlace
a. ikat			*	*!		
b. ikati		*!	*	*		
c. ika	*!		*			*
☞ d. ikaʔ			*		*	*

The position-specificity of place neutralization is captured by the universal ranking of ID-OnsPlace >> MaxPlace. ID-OnsPlace requires that an onset consonant have the Place of its input correspondent, while MaxPlace requires that an underlying Place feature have an output correspondent. Assuming that ID-OnsPlace, which preserves underlying place features in the onset only, always outranks the place hierarchy in (14), a place distinction in coda position is neutralized due to the dominance of the markedness constraints in (14) over MaxPlace. Markedness violations in coda position cannot be resolved by deleting the offending coda due to the high ranking of Max, which penalizes deletion, nor can it be resolved by the addition of a final vowel due to the high ranking of Dep, which penalizes epenthesis. Place distinction neutralizes toward [ʔ] since *Phar, which penalizes /ʔ/, among other things, is ranked lowest relative to the other place markedness constraints; the candidate with a /ʔ/ coda (15d) is thus preferred over the fully faithful candidate (15a), which has a coronal coda.

3.2 Richness of Cues and Contrast Maintenance

As the last case study illustrates, the notion of markedness is often invoked to account for the directionality of contrast reduction. Laryngeal neutralization in coda position is said to favor voicelessness because laryngeal features such as [voice] and [constricted glottis] are more marked than voicelessness (Lombardi 1991). Similarly, the fact that vowels in unstressed position often neutralize to schwa is often attributed to the unmarkedness of schwa. The definition of markedness is a matter of great debate, however. Features and segments that are considered unmarked in some languages or contexts are marked in others. As alluded to earlier, Prince and Smolensky (1993) and Smolensky (1993) propose a place-of-articulation markedness hierarchy, which asserts that coronals are less marked than labials and dorsals, based on the observation that many languages show neutralization of place of articulation in word-final or preconsonantal positions to coronal (see Paradis and Prunet 1991 and Blevins 2004 for more discussion on the unmarked status of coronal). Yet, place neutralization toward labials and velars is also found. Nimboran, a Trans-New Guinea language of North Papua, contrasts labial, alveolar, and velar plosives in word-initial and medial positions (*peb* 'peel', *tebuá* 'short', *kení* 'hearing'; Anceaux 1965: 15–27); in word-final position, only labials are found (e.g. *sip* 'blunt', *sib* 'place'). In Fuzhou (Min Chinese), on the other hand, the only codas allowed are /ŋ/ and /ʔ/. For more discussion of language-specific markedness, see Hume (2003).

In an effort to provide an objective basis for markedness, some scholars have proposed grounding the notion of markedness in terms of the speakers' partial understanding of the physical conditions under which speech is produced and perceived. This phonetically-based notion of markedness leads to the development of a cue-based approach to contrast reduction (see Hayes *et al.* 2004 and references therein). The basic assumption of cue-based approaches to contrast reduction is that a contrast is suspended in positions where the relevant contrast-supporting cues are diminished; a contrast in such cue-impoverished environments may be

maintained only at the cost of additional articulatory maneuvers. A contrast is licensed in positions that are rich in perceptual cues that maximize the contrast's perceptibility. Alveolars and retroflexes, for example, are most distinguishable by their VC transition profiles. Positions where VC transition is impoverished or non-existent, such as word-initial and post-consonantal positions, tend to be loci where the alveolar~retroflex contrast is eliminated (Steriade 1994). For example, Bunuba, an Australian language, contrasts apical alveolar and retroflex word-medially (e.g. *biɖi* 'thigh' vs. *wiɗigi* 'stick insect'), but only apical alveolars are found word-initially (Rumsey 2000). The only exception to this restriction is when a subsequent syllable contains [ɖ, ɳ, ɭ]; in such instances (e.g. *ŋaɖʌ* 'short', *ɖuɭu* 'heart'), long-distance retroflexion is assumed to be what licenses the presence of retroflexion word-initially (Hamann 2003). Even when VC transitions are present, however, retroflexes are often avoided in the environment of /i/. For example, retroflex fricative and affricate series in several Chinese dialects are in complementary distribution with the alveo-palatals: before a high front vowel, only alveo-palatals are found while the retroflexes occur elsewhere (Yip 1996). Hamann (2003) explains this avoidance of retroflexes in the environment /i/ as a result of the articulatory incompatibility between the production of these segments; a flat tongue middle and retracted tongue back configuration for retroflexion cannot be combined with the high tongue middle and fronted tongue back necessary for front vowels. Languages often restrict the distribution of contour tones to phonemic long vowels (e.g. Somali and Navajo), stressed syllables (e.g. Xhosa and Jemez), and word-final positions (Zhang 2001, 2002). While it is difficult to characterize these positions in structural or prosodic terms in a unifying way, they have in common rhyme durations that are long, sonorous, and high in intensity. This fact has led some researchers to hypothesize a long sonorous rhyme duration as the unifying factor for privileged contour tone licensers (Gordon 1999, 2001; Zhang 2001). Obstruents are often voiced after a nasal, resulting in voicing neutralization (Luyia (Niger-Congo) /N + p, t, k, ts, c/ → [mb, nd, ŋg, nz, ñj]; Herbert 1986: 236). Hayes and Stivers (2000) attribute the preference for post-nasal voicing to the effects of "velar pumping" which arises from vertical motion of a closed velum, and of "nasal leak", the leakage of air through a nearly closed velar port during the coarticuatory period between oral and nasal segments.

Structure-based accounts have difficulties accounting for languages, such as Lithuanian, which licenses laryngeal contrasts in pre-sonorant position, regardless of whether the following sonorant is tautosyllabic or heterosyllabic (see also Ancient Greek and Sanskrit; Steriade 1999b). That is, assuming obstruent+sonorant sequences are syllabified heterosyllabically, the fact that obstruents neutralize in voicing in some coda positions but not others is not expected within a structure-based account, which cannot differentiate the different coda positions. From a cue-based perspective, the reduction of laryngeal contrasts in preconsonantal and final positions follows from the fact that many of the relevant cues for the perception of voicing (closure voicing, closure duration, duration of preceding vowel, F_0 and F1 values in preceding and following vowels, VOT values, burst duration and amplitude) are endangered in those positions. The more

impoverished the available perceptual cues are, the less sustainable the laryngeal contrast is. Thus, word-initial preconsonantal position is least hospitable to a contrast in voicing while inter-sonorant position is most ideal for voicing realization. Formally, a cue-based account of contrast reduction may be modelled as the interaction between constraints on contrast maintenance and markedness constraints induced from phonetic knowledge (Hayes 1999; Steriade 1999b). Steriade (1999b), for example, models [voice] neutralization in terms of the interaction between the constraint, Preserve [voice], which demands faithfulness to input voice values, and a fixed hierarchy of *voice constraints aligned to a voice perceptibility scale (16).

(16) Scale of obstruent voicing perceptibility according to context (Steriade 1999b: 11)[8]

V__[+son] ▷ V__# ▷ V__[–son] ▷ [–son]__[–son], [–son]__#, #__[–son]

A language with voicing licensed only before sonorants would have the following ranking:

(17) Voice licensed before sonorants (Steriade 1999b: 12)

*voice/[–son]__[–son], [–son]__#, #__[–son] >> *voice/V__[–son] >> *voice/V__# >> **Preserve[voice]** >> *voice/V__[+son]

Given that the ranking of constraints projected from a phonetically-grounded perceptibility scale has been argued to be universal, such a model makes strong predictions about the typology of laryngeal neutralization patterns. For example, it predicts that a language with a voicing contrast in word-initial preconsonantal position must also allow a voicing contrast in word-initial, intervocalic, and word-final positions. The Mon-Khmer language, Khasi, spoken in the Assam province of India, shows that such a strong prediction does not obtain. As illustrated in (18), Khasi contrasts voiced and voiceless plosives in word-initial preconsonantal position.

(18) Voicing contrast in initial clusters in Khasi (Henderson 1992: 62)

bti	'to lead by the hand'	pdot	'throat'
bthi	'sticky'	pdeng	'middle'
dkar	'tortoise'	tbian	'floor'
dkhar	'plainsman'	tba	'to feel'
dpei	'ashes'	pjah	'cold'
bshad[bʃaːt]	'civet'	bdi	'twenty'

Using evidence from a Frøkjær-Jensen combined oscilloscope and mingograph, Henderson (1992) confirmed the voicing contrast in word-initial preconsonantal position and ruled out the possibility of a *svarabhakti* vowel between the two

stops. What is of interest here is the fact that, in syllable-final position, there is no distinction between voiced and voiceless stops; final stops are unreleased and frequently accompanied by simultaneous glottal constriction (Henderson 1967: 567).[9] Since a voicing contrast is allowed word-initially before another obstruent – a highly impoverished environment for the maintenance of a voicing contrast – a cue-based approach that maintains the universality of voicing perceptibility necessarily predicts that a voicing contrast should also be maintained in less impoverished environments, such as post-vocalic word-final positions. It is worth noting that counterexamples of this sort do not obviate the validity of a cue-based approach to contrast reduction *per se* since the assumption of the universality of cue perceptibility is logically independent of the claim that cue maintenance is the driving force behind contrast maintenance and reduction (see Hume and Johnson 2001 for a recent discussion on the language-specificity of speech perception).

Some cue-based theorists eschew the notion of markedness at the level of the individual segment or feature and favor instead a contrast maximization account. Dubbed "Dispersion Theory" (Flemming 1995, 1996, 2001; Ní Chiosáin and Padgett 2001; Padgett 2003) after Lindblom's (1986; 1990) "Theory of Adaptive Dispersion," such a theory of contrast maintains that the selection of a phonological contrast is subject to three functional goals (see Martinet 1952, 1955, 1964 for early formulations of these functional ideas; cf. Silverman 1996, 2004, 2006):

(19) 1. Maximize the distinctiveness of contrasts.
2. Minimize articulatory effort.
3. Maximize the number of contrasts.

From this perspective, the dispreference for sound *x* is conceptualized as a dispreference for the sub-maximally distinct contrasts between *x* and other sounds in the particular sound system. As schematized in the ranking in (20), a contrast is formally neutralized in some context if it cannot be realized with a distinctiveness of *d* without violating *Effort, an effort minimization constraint penalizing some articulation.

(20) Mindist = *d*, *Effort >> Maximize contrasts

In Belorussian, for example, a five-vowel inventory /i, e, a, o, u/ is observed in stressed syllables. In unstressed syllables, /e, a, o/ reduce to [a] or [ɐ] depending on the position of the vowel relative to the stressed syllable (Barnes 2002: 65). Flemming (2004) argues that this type of vowel reduction is motivated by difficulties in producing distinct F1 contrasts in unstressed positions. Specifically, increasing difficulty in producing a low vowel as a result of vowel duration shortening in unstressed positions leads to the raising of short low vowels; the smaller range of the F1 dimensions for distinguishing F1 contrast then leads to the selection of a smaller number of contrasts. Flemming captures this intuition in terms of the ranking in (21).

(21) Unstressed vowels are short, *Short low V, Mindist=F1:3 >> Maximize contrasts >> Mindist=F1:4

The constraint, Unstressed vowels are short, requires unstressed vowels to be shorter than stressed ones. This constraint will be left out of the subsequent discussion since this is assumed to be undominated and no vowel systems violating this constraint will be admitted in the present context. *Short low V (abbrev. *a) is an effort-minimization constraint that penalizes low vowels. The Mindist=Y:X constraints are satisfied by contrasting sounds that differ by at least X distance on the Y dimension. The highest ranking Mindist constraint that outranks the Maximize contrasts constraint sets the threshold distance, and the optimal inventory is the one that packs the most contrasting vowels onto the relevant dimension (here, F1) without any pair being closer than this threshold.

With the relative positioning of vowels on the F1 dimension stated in (22), Belorussian's three-way vowel height distinction in stressed syllables is predicted in (23). Since the present evaluation concerns only distinctions in vowel height, the back counterparts of vowels in the inventory candidate set are left out for ease of reference. The tableau in (23) shows that a four-way height distinction is suboptimal (23c) because vowels are not distinct enough according to the constraint, Mindist=F1:3. Reducing the height inventory too much (23a) results in excessive contrast reduction, thus incurring more Maximize contrasts violations relative to the optimal inventory set (23b).

(22)

F1:	7	6	5	4	3	2	1
	a	ɐ	ɛ ɐ	e ə	ɐ̟	ɪ	i

(23) Belorussian: vowels in stressed syllables

	*a	Mindist=F1:3	Max contrasts	Mindist=F1:4
a. í-á			✓✓!	
☞ b. í-é-á			✓✓✓	**
c. í-é-é-á		*!	✓✓✓✓	***

In the unstressed syllables, the constraint *Short low V (*a) becomes applicable. It rules out the candidate vowel inventory [i-e-a] because of the presence of [a]. The three-way height distinction cannot be maintained even if the low vowel [a] is avoided, the distance between [e] and [ɐ] being insufficient due to the high ranking Mindist=F1:3 constraint. The winning candidate has only two vowel heights, which fares worse by Max contrasts, but satisfies the higher-ranked minimum distance requirements.

(24) Belorussian: vowels in unstressed syllables

	*a	MINDIST=F1:3	MINDIST=F1:4	MAX CONTRASTS
☞ a. i-ɐ				✓✓
b. i-e-ɐ		*!	**	✓✓✓
c. i-e-a	*!		**	✓✓✓

Within Dispersion Theory, the objects of analysis are systems of oppositions. The notion of contrast reduction is thus given a genuine expression in such an analysis. Whereas most other approaches view mergers and neutralization as the results of the application of constraints or rules that prevent the expression of individual segments or features, Dispersion Theory holds that mergers and neutralization follow from the number of oppositions a language makes available in different contexts. It should be noted that, because of its insistence of looking at systems of contrast from the perspective of the language as a whole, Dispersion Theory raises questions about how phonological derivation is implemented in such a model (Boersma 1998: 361; but see Ní Chiosáin and Padgett 2001 and Padgett 2003 for a response to this problem).

This section has reviewed major theories of contrast reduction, showing that proposals range from completely structure-dependent accounts to theories that embrace the full phonetic substance of sound patterns. The debate on what a proper theory of contrast reduction is, however, might ultimately rest on resolving a more fundamental question – does synchronic contrast reduction truly exist? This is the topic of the next section.

4 Do Real Synchronic Mergers and Neutralization Exist?

Until recently, most theories of phonology assume some form of lexical minimality (the minimization of lexically stored information; Chomsky and Halle 1968: 381, Steriade 1995: 114) and feature economy (the minimization of the ratio of features to segments in an "alphabet"; Clements 2003). In early generative phonology, for example, the underlying alphabet is the minimal sound set needed to express surface differences between distinct morphemes; at the level of the underlying representation, no allophonic variants are present. Theories differ in the number of levels of representation allowed (e.g. Lexical Phonology and Morphology (Kiparsky 1982, 1985; Mohanan 1982) recognizes three levels of representations: underlying, lexical, and phonetic and the degree of minimality assumed at each level. Common across these early theories of phonology, however, is the premise that, out of the vast sea of phonetic signals, only a small subset of phonetic properties are contrastive in a given language (Sapir 1949; Trubetzkoy 1939; Jakobson *et al.* 1952;

Hockett 1955; Chomsky and Halle 1968; Kiparsky 1982, 1985). Contrast is encoded in terms of a difference between $+F_i$ and $-F_i$ for some finite set of features F_i, and contrast reduction corresponds to the elimination of this difference (i.e. the outcome of such a reduction is either $+F_i$, $-F_i$, or null).

Non-distinctive phonetic properties are treated in one of two ways. To begin with, features that do not distinguish lexical items may be underspecified in the lexical entries (e.g. Archangeli 1988; Pulleyblank 1995; Steriade 1995; Clements 2003). The feature [voice] in sonorants, for example, is non-contrastive, and thus redundant, in languages such as English, which do not distinguish between voiced and voiceless sonorants. Sonorants are underspecified for voicing, that is, sonorants bear no value for the feature [voice]. Such an assumption of non-contrastive feature underspecification has important theoretical consequences for the treatment of transparency effects in the phonology of the feature [voice]. For example, as seen earlier, nasals do not induce regressive voicing assimilation in Dutch, but voiced obstruents do, suggesting that only voiced obstruents are underlyingly specified for the feature [voice]. The other treatment of non-distinctive phonetic properties is to exclude them from the feature pool altogether. For example, vowels are longer before voiced stops than before voiceless ones in American English (*bat* [bæt] vs. *bad* [bæ'd]). Peterson and Lehiste (1960) suggest that the ratio of vocoid duration before voiceless consonants to that before voiced consonants in American English is 2:3. Such a difference in vocoid duration which covaries with the voicing of the following consonant is generally dismissed as the effect of automatic phonetics, and thus assumed to play no role in any phonological analysis; features such as [slightly long] would not be part of the universe of phonological features.[10]

As Labov *et al.* (1991) point out, the assumptions that "contrasts were discrete and binary, that there was no such thing as a small difference in sound, that production and perception were symmetrical, and that introspections were reliable" (p. 38) have received increased scrutiny in recent years. For example, Dispersion Theory's admission of phonological constraints that regulate features along scalar dimensions, rather than in terms of binary oppositions, already foreshadows the move away from a discrete and binary notion of contrasts (e.g. MINDIST=Y:X constraints evaluate distances along some phonetic dimensions such as F1). Mounting evidence for near mergers and incomplete neutralization raises further questions about the validity of these above-mentioned assumptions. This is the topic of the next section.

4.1 Near Mergers and Incomplete Neutralization

Near merger describes the situation where speakers consistently report that two classes of sounds are the same, yet consistently differentiate them in production at a better than chance level. Labov *et al.* (1972: Chapter 6), for example, reports that speakers in New York City differentiate words such as *source* and *sauce* in production, but report no distinction between them in perception. Similar near mergers have been reported in other varieties of English (e.g. *fool* and *full* in

Albuquerque (Di Paolo 1988); *too* vs. *toe* and *beer* vs. *bear* in Norwich (Trudgill 1974); *line* vs. *loin* in Essex (Labov 1971; Nunberg 1980); *meat* vs. *mate* in Belfast (Milroy and Harris 1980; Harris 1985). Near mergers do not only concern segmental contrasts. Yu (2007b), for example, showed that derived mid-rising tones in Cantonese show a small but statistically significant difference in f_0 from underived mid-rising tones. Similar to near mergers, *incomplete neutralization* refers to reports of small but consistent phonetic differences between segments that are supposedly neutralized in certain environments. Flapping is often cited as a neutralizing phonological alternation in American English; underlying /t/ and /d/ surface as dental flaps or taps when followed by an unstressed vowel.[11] For example, when the final consonants in [siːt] 'seat' and [siːd] 'seed' appear intervocalically, as in 'seater' and 'seeder', both underlying /t/ and /d/ surface as either dental flaps or taps ([siːɾə˞]). Researchers have long observed that flaps derived from underlying /t/ are not completely identical to the flaps derived from underlying /d/, both in terms of the realization of the flap itself and also in the effect of the flap on its neighboring segments. Trager (1942) described the segment produced by the flapping of /t/ as 'voice *t*, in opposition to a *d*.' Fox and Terbeek (1977) found a significant durational difference between vowels preceding flaps with underlying /d/ versus flaps with underlying /t/. They also found that the first vowel/second vowel ratios are significantly higher in words with /d/-flaps than those with /t/-flaps.

Word-final and preconsonantal obstruent devoicing is another classic example of a neutralizing sound pattern. German, for example, contrasts between consonants such as /t/ and /d/. While this contrast occurs word-initially (e.g. *Tier* [tiːr] 'animal' vs. *dir* [diːr] 'to you') and intervocalically (e.g. *leiten* [laitən] 'lead' vs. *leiden* [laidən] 'suffer'), only voiceless consonants are found word-finally. Production and perceptual experimental results, however, show that the two sets of final voiceless stops are consistently different phonetically (Port and O'Dell 1985; Port and Crawford 1989). For example, compared to words like *Rat* /rat/ 'advice', words like *Rad* /rad/ 'wheel' have longer vowel duration and their final consonants have longer stop closure voicing duration and shorter stop closure and burst duration. Dinnsen and Charles-Luce (1984) found that, in Catalan, underlying voice distinction of word-final obstruents is preserved in either the contextual shortening of the preceding vowel or in the overall closure duration of the final obstruent. For example, in Catalan, vowels shorten before word-final obstruents if the next word begins with a consonant. For some speakers, however, there is a difference in the relative shortening of the vowel, depending on whether the following word-final obstruent is underlyingly voiced or voiceless; vowel shortening is significantly less if the following obstruent is underlyingly voiced. Other speakers preserve the underlying voice distinction in the closure duration of the final obstruent; underlying voiced obstruents are longer than underlying voiceless ones. Similar findings of incomplete neutralization of voicing in word-final and pre-consonantal positions have been reported for other languages, including Russian (Chen 1970; Burton and Robblee 1997), Polish (Giannini and Cinque 1978; Tieszen 1997), Lezgian (Yu 2004), and Dutch (Warner *et al.* 2004).

Incomplete neutralization has been reported outside the domain of obstruent voicing as well. In Eastern Andalusian Spanish, for example, the combined effect of word-internal coda aspiration and the gemination of the consonant following the aspirated coda leads to potential neutralization (e.g. [ka^{h}tːa] for both /kasta/ 'caste' and /kapta/ 's/he captures'). Gerfen (2002), however, reports that aspirating an /s/ results in a longer duration of aspiration, while aspirating a /p/ or /k/ results in longer medial consonant gemination (see also Gerfen and Hall 2001). Bishop (2007) found that listeners make use of the length of the consonant following aspiration as a cue for making phonemic decisions regarding the nature of the underlying coda. In many languages, an epenthetic stop can occur within nasal-fricative clusters or heterorganic nasal-stop clusters (e.g. English 'dreamt' [dɹɛmt]~[dɹɛmpt]; 'prince' [pɹɪns]~[pɹɪnts]). Several studies have found that such epenthetic stops are phonetically different from underlying stops that occur in the same environment. Fourakis and Port (1986), for example, found that underlying /t/ in words like 'prints' [pɹɪnts] are significantly longer in duration and the neighboring nasal significantly shorter than epenthetic [t] in words like 'prince'. Dinnsen (1985), citing Rudin (1980), reports that long vowels deriving from underlying /VgV/ sequences in Turkish are 13% longer than underlying long vowel /Vː/. Simonet *et al.* (2008) report that the so-called /r/ and /l/ neutralization in post-nuclear position in Puerto Rican Spanish (e.g. /árma/→[álma] 'weapon' vs. /álma/→[álma] 'soul') is incomplete. Based on measurements of duration of the vowel+liquid sequences and examination of formant values and trajectories, Simonet *et al.* (2008) conclude that, while post-nuclear /r/s are similar to post-nuclear /l/, there nonetheless exist systematic durational and spectral differences, suggesting that the two liquids have not completely merged.

Since early theories of the phonetics-phonology interface (Cohn 1990; Keating 1990; Pierrehumbert 1990) assume that phonological representations in the lexicon are categorical, contrastive elements, and since the phonetic implementation component computes the degree and timing of articulatory gestures, which are gradient and variable, the discovery of near mergers and incomplete neutralization presents a curious conundrum. For a given underlying distinction +F and –F, how can an output –F that corresponds to an underlying +F display systematically different surface phonetic realization from an output –F that corresponds to an underlying –F, when information flow is supposed to be strictly unidirectional? That is, in such a model, no articulatory plan can look backward to phonological encoding, nor can phonological encoding look back to the lexical level. No lexical information can influence the phonetic implementation directly either, bypassing the level of phonological encoding. On this view, the categorical form of a lexeme determines the general phonetic outcome. Phonetic variations on the surface are considered artifacts of the context or performance-induced anomalies.

In light of such conceptual difficulties, many have sought to explain away the observed subphonemic phonetic differences as a consequence of orthographic influence or as variation in speaking style. For example, it has been found that the less the experimental design emphasizes the role of orthography, the smaller the durational effects (Fourakis and Iverson 1984; Jassem and Richter 1989). Port

and Crawford (1989) found that discriminant analysis to classify productions by underlying final voicing was most successful (78% correct) when speakers dictated the words, but least successful (approximately 55% correct) when target words were embedded in sentences that do not draw attention to the minimal pairs (whether read or repeated orally). But attributing all cases of near mergers and incomplete neutralization to performance factors is insufficient. Warner *et al.* (2004), for example, found subphonemic durational differences in the case of final devoicing in Dutch, even when possible orthographic influence was controlled for as a confound, although they later showed that there is an orthographic component to incomplete neutralization (Warner *et al.* 2006). Yu (2007b) found incomplete merger of underived and morphologically-derived mid-rising tones in Cantonese, a language whose orthography does not indicate tone. Further supporting the existence of a suspended contrast comes from the fact that speakers appear to have some access to subtle phonetic differences. As noted earlier, Bishop (2007) found that Andalusian Spanish speakers can make use of subtle closure duration differences to recover underlying coda consonants. In the case of final devoicing in Dutch, listeners can not only perceive durational differences (Warner *et al.* 2004), they even use these subphonemic distinctions to hypothesize which past tense allomorph nonce forms would take (Ernestus and Baayen 2003).

4.2 Approaches to Subphonemic Phonetic Differences

Subphonemic distinctions have been analyzed as the result of paradigm uniformity among morphologically-related neighbors (e.g. phonetic analogy; Steriade 2000; Yu 2007c). Steriade (2000), for example, argues that grammars prefer words within a paradigm to be uniform.[12] Such an effect of paradigm uniformity is observed when there exists systematic generalization of one allomorph to positions where it is phonologically unjustified or unexpected. Steriade extends this paradigm uniformity preference to the phonetic level. French, for example, has an optional schwa deletion which creates ostensibly homophonous strings (e.g. *bas retrouvé* [baʁətʁuve] 'stocking found again' → *bas r'trouvé* [baʁtʁuve] vs. *bar trouvé* [baʁtʁuve] 'bar found'). Various studies have shown that the consonant to the left of the syllable of the deleted schwa maintains phonetic qualities that would only be expected if the schwa were still present (Fougeron and Steriade 1997; Rialland 1986). Steriade (2000) interprets such unexpected phonetic differences as the results of phonetic analogy; forms with schwa deletion are influenced phonetically by the corresponding schwa-ful forms (e.g. /r/ in *bas r'trouvé* [baʁtʁuve] takes on onset-like articulation from the /r/ in the related phrase *bas retrouvé* [baʁətʁuve]).

Van Oostendorp (2008) argues that incomplete neutralization in final devoicing can be captured within a Containment model of OT in terms of a turbid representation of phonological outputs (Goldrick 2001). Output structures are characterized in terms of two types of relations: a *Projection* relation, which is an abstract structural relationship holding between a segment and the feature (represented by ↑ in (25)), and a *Pronunciation* relation, an output relationship that holds between

the feature and the segment and describes the output realization of a structure (represented by ↓ in (25)). Within this conception, a three-way distinction obtains between segments that are underlyingly voiceless (i.e. they lack the feature [voice]), segments that are underlyingly voiced and pronounced voiced, and segments that are underlyingly voiced but are not realized as voiced on the surface (25).

(25) A three-way voicing distinction using turbidity theory

a. ta:[t]

b. ta[d]
↑↓
[voice]

c. ta:[d̥]
↑
[voice]

The selection of a representation like (25c) would be determined by the interaction between markedness constraints that disfavor coda voicing and the constraint, Reciprocity(X, F), which holds that if a segment X entertains a *projection* relation with a feature F, then F must entertain a *pronunciation* relation with the segment X. Because of their structural differences, (25)a, b, and c will show different surface phonetic realization.

These phonological approaches assume that cases of incomplete neutralization are in fact complete at the phonological level. That is, they assume that the output segment is phonologically unvoiced. The subphonemic differences observed would either be due to analogical influences from related forms that retain voicing or to covert structural differences among outputs. Is a complete neutralization interpretation of incomplete neutralization a necessity, or even desirable? The answer to this question hinges on the conception of the phonetics-phonology interface and, specifically, the nature of allophony. Which allophones should be considered extrinsic (i.e. phonologically governed), and which should be considered intrinsic (i.e. introduced by phonetic variability; Wang and Fillmore 1961; Ladefoged 1971; Tatham 1971)? Must extrinsic allophones be governed by changes in discrete distinctive feature values or can extrinsic allophones be gradient? The next section offers an alternative interpretation of near mergers and incomplete neutralization, appealing to the notion of a covert contrast.

4.3 Subphonemic Distinctions as Covert Contrasts

Near mergers and incomplete neutralization are problematic from the point of view of a model of the interface between phonetics and phonology sketched above because, if the phonetic implementation component accounts only for variations due to biomechanical and aerodynamic factors, it is anomalous, to say the least, that speakers of a language with [voice] neutralization vary the realization of the neutralized sounds in accordance with the feature value of their non-neutralized counterparts. The above model of the phonetics-phonology interface is arguably simplistic, however. Kingston and Diehl (1994) articulate a model of the phonetics-phonology interface that affords the phonological component greater control over

the range of variability in the phonetic implementation of contrasts. Elasto-inertial, biomechanical, aerodynamic, psychoacoustic, and perceptual constraints delimit what a speaker (or listener) *can* do, but not what they *must* do. Within this conception of the phonetics-phonology interface, a phonemic contrast is taken to be "any difference in the feature content or arrangement of an utterance's phonological representation which may convey a difference in semantic interpretation" and allophones are "any phonetic variant of a distinctive feature specification or arrangement of such specification that occurs in a particular context" (p. 420 note 2). To illustrate this framework more concretely, consider Kingston and Diehl's (1994) summary of the phonetic variants of English stops contrasting for [voice] (see also Silverman 2004):

(26) Summary of the phonetic variants of English stops that contrast for [voice].

	[+voice]	[−voice]
Utterance-initial or pretonic	short lag VOT	long lag VOT
	F_1 lower	F_1 higher
	F_0 lower	F_0 higher
	weaker burst	stronger burst
Intervocalic or posttonic	closure voicing	no closure voicing
	short closure	longer closure
	longer preceding vowel	shorter preceding vowel
	F_1 lower	F_1 higher
	F_0 lower	F_0 higher
Utterance-final and postvocalic	longer preceding vowel	shorter preceding vowel
	closure voicing possible	no closure voicing
	short closure	longer closure
	F_1 lower	F_1 higher

(26) illustrates the fact that the contrastive feature [+voice] in English shows great variability in its phonetic realization. In word-initial position, for example, [+voice] stops are often realized as voiceless unaspirated, even when the preceding word end in a vowel (Caisse 1982; Docherty 1989). Kingston and Diehl (1994) interpret such data as showing that speakers choose between two active articulations in producing initial [+voice] stops in English: delay glottal closure until the stop release, or close the glottis but expand the oral cavity to overcome the difficulty of initiating voicing. Such controlled variation is made possible by the fact that there are typically multiple, auditorily independent correlates that serve as distinct bases for a minimal phonological distinction. As noted in Stevens and Blumstein (1981), [+voice] consonants are characterized by the "presence of low-frequency spectral energy or periodicity over a time interval of 20–30 msec in the vicinity of the acoustic discontinuity that precedes or follows the consonantal constriction interval" (1981: 29). This low-frequency property, as Kingston and Diehl (1994)

call it, has multiple supporting subproperties such as voicing during the consonant constriction interval, a low F1 near the constriction interval, and a low F_0 in the same region, as well as enhancing properties such as the duration ratio between a consonant and its preceding vowel. These properties do not all surface in all positions. Crucially, while [+voice] stops do not show prevoicing in word-initial position, the [voice] contrast is nonetheless maintained because [–voice] stops tend to have longer VOT, stronger burst energy, and higher F1 and F_0 following the consonant constriction interval.

From this perspective of the phonetics-phonology interface, subphonemic differences observed in near mergers and incomplete neutralization are qualitatively not different from those observed between allophones appearing in different phonetic contexts. As noted in Steriade 1999b, the perception of voicing hinges on a multitude of acoustic cues: burst amplitude, closure duration, voicing during the closure period, voice onset time, and vowel onset and offset. Phonetic cues that support a [voice] contrast in word-final positions are intrinsically impoverished relative to cues available in word-initial and word-medial positions. Nonetheless, many languages maintain the contrast in word-final positions because there remain sufficient cues that can differentiate the underlying phonological contrast. To illustrate this approach, let us adopt Flemming's scalar approach to contrast maintenance, which conceptualizes contrastivity as a matter of the distance separating contrastive elements along some phonetic dimensions (the supportive subproperties; Kingston and Diehl 1994). (27), for example, shows a hypothetical language which maintains a full [voice] contrast in word-initial and word-medial positions (i.e. non-final); in such positions, [–voice] (i.e. T) and [+voice] (i.e. D) units are maximally separated from each other along the phonetic dimensions given.[13] In word-final positions, certain phonetic cues are less effective at signaling the contrast (e.g. C-voi) or are not relevant at all (e.g. VOT). Crucially, the contrast between [+voice] and [–voice] is maintained nonetheless since [+voice] and [–voice] remain distinct along other dimensions.

(27)

	Non-final								Final							
	7	6	5	4	3	2	1	Dist.	7	6	5	4	3	2	1	Dist.
VOT	D						T	6								NA
C-voi	D						T	6			D				T	4
C-dur	D						T	6	D						T	6
V-dur	D						T	6	D						T	6

In a language that shows incomplete neutralization of final devoicing, the range of phonetic cues that signal the [voice] contrast in coda position might be so impoverished that the contrast may escape detection by traditional methods of linguistic data collection (i.e. impressionistic transcription). Nonetheless, the contrast is maintained from the perspective of the native speaker, albeit covertly (i.e. not available to explicit introspection).

(28) Incomplete neutralization in word-final positions

	7	6	5	4	3	2	1	Dist.
VOT								NA
C-voi						D	T	1
C-dur			D				T	4
V-dur		D					T	5

In a similar vein, Kirby (2010, to appear) argued that, in the case of incomplete neutralization, category separation (i.e. contrast maintenance) may obtain even when individual supporting cues, when considered in isolation, do not warrant such a separation. Kirby (2010, to appear) followed recent work on applying computational techniques of clustering and pattern recognition to the problem of phonological category induction (de Boer and Kuhl 2003; Lin 2005; Vallabha *et al.* 2007; Feldman *et al.* 2009; McMurray *et al.* 2009) and modeled phonetic categories as Gaussian mixtures of cue distributions. Adopting a model-based clustering approach, Kirby demonstrated that a two-category solution (e.g. the retention of a voiced vs. voiceless contrast in word-final position in Dutch) may be justified when multipled cues are considered simultaneously, even when the relevant phonetic cues do not individually support a categorization solution with more than one category (i.e. neutralization). He further showed that the two-category solution is only feasible under certain combinations of cues; other cue combinations would lead to complete merger.

Flapping of /t/ and /d/ in English can be seen as an instance of *covert contrast* at work. As alluded to earlier, many scholars have suspected that flapping in American English might be incompletely neutralized (Trager 1942; Fox and Terbeek 1977). Many have also noted that there is great variability in the output of the flapping rule (Haugen 1938; Zue and Lafferriere 1979; Stone and Hamlet 1982; Turk 1992; de Jong 1998). The flapped consonants may be described variously as alveolar stops, alveolar flaps, or taps. Haugen, for example, characterizes the flapped [t]'s as falling along a continuum from a "fully articulated *d*" to a lightly articulated tap. In his X-ray microbeam study of English speakers producing utterances with word-final coronal consonants in the appropriate segmental context for flapping, de Jong (1998) found support for a prosodically motivated articulatory explanation for flapping, arguing that flapping may be seen as a by-product of consonant-vowel coarticulation and the encoding of prosodic organization in the jaw movement profile (Fujimura 1986; de Jong 1998; de Jong *et al.* 1993). If flapping arises as a by-product of articulatory changes associated with the general implementation of prosodic structure rather than as a categorical switch from a stop to flap production, it suggests that the flapping rule in American English might be an epiphenomenon created by linguists' bias towards categorical analysis. Under this interpretation, the fact that the vowel preceding a flapped /t/ is shorter than that preceding a flapped /d/ (Fox and Terbeek 1977) follows straightforwardly from the fact that vowels preceding voiced obstruents

are longer than vowels preceding voiceless obstruents in general (Peterson and Lehiste 1960).

The interpretation of near mergers and incomplete neutralization advocated here suggests that traditional methods of introspection and field elicitation may not be adequate in detecting covert contrast. Manaster-Ramer (1996) remarked that several giants in phonetics and historical linguistics have noted many cases of neutralization and that it would be surprising if they are all "wrong." Such sentiments cannot be sustained, however. Self-introspection faces inherent problems of analyst bias and thus should not be taken as a definitive source of information. The phonetician's ears are, after all, human ears. Commutation tests are essentially armchair psycholinguistic tasks that require language consultants to perform a same-different task with minimal control for potential confounds. Subject responses are inherently probabilistic; analysts insisting on dichotomizing a continuous function will find confident responses when the samples have a wide separation in the sample space. Samples that straddle regions of great overlap, as in the case of near mergers and incomplete neutralization, will elicit more ambiguous responses. Contrasts not detected by linguists using traditional methods of elicitation may nonetheless be detected by native speakers, as demonstrated in laboratory studies alluded to above (i.e. Ernestus and Baayen 2003; Warner *et al.* 2004; Bishop 2007).

4.4 Covert Contrasts as Systems in Transition

The existence of covert contrast is readily understandable from the perspective of sound change and phonologization of phonetic variation. In his seminal work on phonologization, Hyman (1976) conceptualized the emergence of phonemic tonal distinctions as a three-stage process. At Stage 1, a language displays physiologically-based consonantal voicing-induced pitch perturbations on the neighboring vowel (Hombert 1978; Hombert *et al.* 1979). A language reaches Stage 2 when pitch perturbation becomes exaggerated to such an extent that the pitch variation cannot be attributed entirely to the physiological properties of the preceding consonant's voicing (e.g. *pa > pá and *ba > bà). The transition from Stage 1 to Stage 2 – when an intrinsic, thus unintended, variation in pitch associated with consonantal realization becomes an extrinsic feature of the vowel – is phonologization. A language reaches Stage 3 when the voicing distinction is lost completely and the pitch distinction on vowels becomes the sole feature that signals a meaning difference between words. That is, the language has undergone the phonemicization of tone (i.e. *pa > pá and *ba > pà). From the perspective of this model of sound change, covert contrast represents a language at Stage 2 and possibly in transition to Stage 3. That is, the old contrast (e.g. obstruent voicing) has not completely disappeared (i.e. been neutralized), but the new contrast (e.g. tonal distinction) has not fully emerged either. A language in Stage 2 is in principle unstable. As Hyman points out, "accompanying every phonologization is a potential dephonologization" (Hyman 1976: 410). The emergence of a tonal distinction as a result of the phonologization of intrinsic pitch perturbation of obstruent

voicing entails the eventual destruction (i.e. neutralization) of the original voicing contrast.

The evolution of vowel duration and consonant voicing co-variation provides an instructive example of phonologization and its connection to the emergence of covert contrast. As reviewed in Solé (2007), languages differ in the amount of control the speakers have over the maintenance of this subphonemic duration difference. Solé (2007) found that English speakers actively maintain durational differences before voiced and voiceless stops regardless of speaking rates, while speakers of Catalan and Arabic do not exhibit similar control over such subphonemic duration differences. Her findings suggest that English has already partially phonologized the effect of consonant voicing on vowel duration, while Catalan and Arabic have not. Recall that one commonly observed feature of the incomplete neutralization of final devoicing is a vowel duration difference. Following Hyman's dictum that the phonologization of one feature carries the seed for the destruction of another, the phonologization of a subphonemic vowel duration difference entails an eventual loss of the voicing contrast in the following stop. The reasons why such a correlation exists are still a matter of debate. Two factors are noteworthy in this context. First, the longer vowel before voiced stops and the shorter vowel before voiceless stops are, strictly speaking, in complementary distribution. Likewise, post-vocalic voiced and voiceless stops are also in complementary distribution since they do not appear in the same context. This type of analytic ambiguity (i.e. between vowel duration and consonantal voicing) is typical of a language undergoing phonologization. Second, research on auditory category learning has shown that listeners not only are sensitive to the distributional information of the category cues, they also prefer unidimensional contrasts over multidimensional ones (Goudbeek 2006; Goudbeek *et al.* 2008; Clayards 2008; Clayards *et al.* 2008). Such results suggest that, all else being equal, listeners may rely more heavily on a single cue for category identification even when multiple cues are available in the signal. So, for example, as voicing during stop closure becomes less and less prominent as a feature of voiced stops in final position, vowel length becomes the more reliable contrastive feature. When voicing in closure ceases to be a feature of final obstruents altogether, a contrast in vowel length is expected to emerge. Friulian provides an instructive example of this type of cue trade-off in phonologization. In Friulian, a Romance language spoken in Northeastern Italy, vowel length is only distinctive in stressed word-final syllables closed by a single consonant (29).

(29) Vowel length distinction in Friulian (Baroni and Vanelli 2000: 16)

a.	['laːt]	'gone(m.)'	b.	['lat]	'milk'
	['bruːt]	'brother, mother-in-law'		['brut]	'ugly'
	[fi'niːt]	'finished(m.)'		['frit]	'fried(m.)'
	['paːs]	'peace'		['pas]	'step'
	['fuːk]	'fire'		['tɔk]	'piece'

Stressed vowels are always phonologically long before [r] (['laːrk] 'large(m.)') and always short when they are not in the last syllable of a word ([kan'tade] 'sung(f.)'), when they occur in the final open syllable ([ku'si] 'so'), and when they are in the final syllable closed by a consonant cluster, nasals, or affricates (['ɡust] 'taste', ['maŋ] 'hand', ['bratʃ] 'arm'). Of particular relevance here is the fact that vowel length in word-final syllables before obstruents is predictable: the stressed vowel is long if the following consonant is realized as voiced in intervocalic position (30a); if the following consonant is voiceless intervocalically, the stressed vowel is short (30b).

(30) Vowel length and consonant voicing (Baroni and Vanelli 2000: 17)

a.	['laːt]/['lade]	'gone(m.)/(f.)'
	[fi'niːt]/[fi'nide]	'finished(m.)/(f.)'
	['peːs]/[pe'za]	'snow/to snow'
	['fuːk]/[fogo'laːr]	'fire/fireplace'
b.	['lat]/[la'te]	'milk/to breast-feed'
	['pas]/[pa'sa]	'pass/to pass'
	[pa'taf]/[pata'fa]	'slap/to slap'
	['tɔk]/[tu'kut]	'piece/little piece'

Based on acoustic evidence, Baroni and Vanelli (2000) establish that long vowels are more than twice as long as short ones and word-final obstruents are indeed voiceless (i.e. no voicing during closure). While final -[t]s that correspond to medial [d] are significantly shorter than final -[t]s that correspond to medial [t], such a difference is only observed after certain vowel qualities. Their findings suggest that, while Friulian final obstruent devoicing is incomplete (i.e. there remains some difference between underlying /d/ and underlying /t/ in final positions), this difference is mainly carried by the closure duration of the obstruent and only in very restricted contexts. On the other hand, a full-blown vowel quantity difference has emerged in its place. The salience of this vowel length contrast is exemplified by the behavior of vowel length in loanword adaptation. Friulian does not preserve the consonantal length contrast in borrowed Italian words. However, longer vowels before single consonants in Italian are treated as long in Friulian if they occur in word-final position: [impje'gaːt] 'clerk(m.)', from Italian [impje'gaːto]. When such long vowels occur in word-internal position, they become short and the following obstruent is voiced ([impje'gade] 'clerk(f.)'). When borrowed short vowels occur in word-internal position, the obstruents remain voiceless (e.g. [a'fit]/[afi'tut] 'rent/little rent'; Italian [a'fːitːo]). This loanword evidence suggests that Friulian has restructured its system to one with a limited vowel length contrast; voicing variation has become secondary to vowel length difference.

Kirby (2010, to appear) argued that transphonologization of this sort (i.e. the phonologization of one cue along with the dephonologization of another) can be seen as a natural consequence of adaptative enhancement in speech communication.

Working with the assumption that learners monitor the precision of contrast distinction and the relative informativeness of the acoustic cues signaling, Kirby demonstrated through a series of computational simulations that a learner would show bias toward enhancing the most informative acoustic-phonetic cue to a contrast. For example, in the case of post-vocalic obstruent voicing, if the informativeness of VOT is diminished as a result of some external bias, vowel duration before obstruents would be enhanced if the same lexical contrast is to be maintained. The likelihood of a secondary cue being enhanced (vowel duration in the case of Friulian) is contingent on its informativeness relative to other relevant cues (e.g. closure duration, f_0 before closure, etc.) and the precision of the contrast. Many factors (e.g. speaker prior experience, attention, etc.) potentially influence the relative informativeness of perceptual cues to a contrast. Nonetheless, the basic framework of understanding transphonologization as a consequence of the cue-trading relationship between the primary and secondary cues offers great potentials for explaining the overall tendencies of (incompletely) neutralizing sound changes.

The possibility of a covert contrast in purported cases of neutralization raises questions about the existence of genuine instances of neutralization. Kim and Jongman (1996), for example, reported that coda neutralization (i.e. word-final coronal obstruents (e.g. /t, t^h, s/ are all phonetically realized as [t]) is complete in Korean. Based on both production and perceptual data, they concluded that complete neutralization is observed despite the fact that Korean orthography distinguishes between the different underlying consonants. The difference between genuine neutralization vs. covert contrasts might also be related to the nature of the evidence supporting the claim of neutralization. Evidence for neutralization may come from distributional information alone (e.g. laryngeal neutralization in coda position in Cantonese) or may additionally be supported by morphological means (e.g. obstruent-devoicing in Dutch). All reported cases of incomplete neutralization pertain to morphologically-sensitive neutralization. In a similar vein, Warner *et al.* (2006) suggested that incomplete neutralization might be restricted to positional neutralization phenomena while global neutralization leads to complete neutralization. In sum, the question of how pervasive covert contrasts and complete neutralization are is ultimately an empirical one. More systematic phonetic and psycholinguistic investigations are needed to answer this fundamental question in contrast reduction research. As contrast reduction has traditionally been a great source of inspiration for research on phonological features and their organizations, the existence and potential pervasiveness of incomplete neutralization and near merger point to a need for reconsidering the nature of phonological features and refocusing attention on the phonetic makeup of phonological contrasts.

ACKNOWLEDGMENTS

This paper is nearly identical to my paper on "Mergers and neutralization" in the *Companion to Phonology* (Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren

Rice, eds., Wiley-Blackwell, 2010) with the exception of Section 4, which has been revised to reflect more closely my current thinking on the nature of phonological contrasts and their reduction. I thank James Kirby, Beth Hume, Colin Ewen, Marc Oostendorp, and two anonymous reviewers for helpful comments on earlier incarnations of this chapter.

NOTES

1 Unless noted otherwise, I shall abstract away from the issue of context sensitivity in what follows.

2 The definition of neutralization adopted here differs from Kiparsky (1976: 169)'s formulation of a neutralizing rule, which states that a rule of the form A → B/XC_DY is neutralizing iff there are strings of the form CBD in the input to the rule. Certain structure-building neutralizing rules, such as flapping in English, are not considered neutralizing from Kiparsky's perspective since the product of the rule is not phonemic in the language.

3 Yu (2004) reports a case of final neutralization toward the voiced series in Lezgian (North Caucasian); the neutralization is restricted only to monosyllabic nouns, however. The default direction of neutralization in final position is toward the voiceless aspirated series.

4 See Silverman (2010) for a thorough review of neutralizing processes in Korean.

5 C′ indicates a tense consonant, not an ejective.

6 Within Optimality Theory, two types of constraints have been posited to account for positional asymmetries in the realization of segmental features. Positional licensing constraints (Itô *et al.* 1995; Walker 2001, 2005; Zoll 1996, 1998), which find analogs in earlier accounts of positional asymmetries (Goldsmith 1989, 1990; Trubetzkoy 1939), assign violations when a marked feature *f* appears unassociated with a licensing context *C* in the output. Positional faithfulness constraints assess violations when an output segment in a privileged context *C* differ in value for the feature *f* relative to its input correspondent (Beckman 1997, 1998; Casali 1996; Lombdardi 1999).

7 The CodaCon constraint, which bans any Place feature in coda consonants, is left out of this table because it is not directly relevant in the present evaluation.

8 The ▷ symbol in (16) indicates that voicing is more perceptible in the context to its left than to the context of its right.

9 While Khasi orthography shows a distinction between *p, t* and *b, d* in syllable-final position (e.g. *pad* 'fathom' vs. *pat* 'then'), the consonant symbols are used to indicate a difference in the length of the preceding vowel; long vowels come before *b* and *d*, and short vowels before *p* and *t* (e.g. *pad* [paːt] 'fathom' vs. *pat* [pat] 'then'). According to Henderson (1967), this spelling convention is ascribable to the link between vowel length and voicing of final consonants in Welsh, since the Khasi orthography was designed by Welsh missionaries.

10 While these sub-featural cues might not be distinctive, they may nonetheless have enhancing functions (Stevens *et al.* 1986; Stevens and Keyser 1989; Keyser and Stevens 2001).

11 In most formulations of this rule, the coronal must occur between a stressed and an unstressed vowel. The preceding stressed vowel is not a requirement, however, as evidenced in the flapping of /t/ in words such as *uniformity, obesity, calamitous, speedometer.*

12 A paradigm is defined here as "a set of words sharing a morpheme, e.g. {*bomb, bomb-ing, bomb-ard, . . .*}, or a set of phrases sharing a word, e.g. {*bomb, the bomb, . . .*}" (Steriade 2000).

13 Only a subset is given here for simplicity's sake; VOT = Voice Onset Time, C-voi = Voicing during consonant closure, C-dur = Closure duration, V-dur = duration of the preceding and/or following vowel.

10 Diachronic Explanations of Sound Patterns

GUNNAR ÓLAFUR HANSSON

1 Introduction: The Phonetic Bases of Phonology

As even a cursory acquaintance with phonology will reveal, the vast majority of sound patterns and phonological processes in the world's languages show clear indications of being shaped by the physical constraints of speech production and perception. Much of phonology is thus arguably "natural" from a phonetic perspective. Distributional patterns and asymmetries in the ability of different positions to support phonemic contrasts (syllable onset vs. coda, stressed vs. unstressed syllable) are strongly correlated with the relative salience of acoustic-perceptual cues and/or the relative magnitude of articulatory gestures in the positions in question. Common phonological processes usually have clear parallels in low-level patterns of phonetic variation, both in the articulatory domain (coarticulation, gestural undershoot, inter-gestural timing and overlap, boundary strengthening) and in the acoustic-perceptual domain (confusion, misperception, aerodynamic effects on the acoustic signal). Such obvious and pervasive correlations raise a number of questions which cut to the conceptual core of phonology as a discipline. What is the nature of the connection between phonetic "substance" and the higher-order patterning that phonology deals with? How close is this connection, and how – if at all – should we allow it to inform our theoretical models of the implicit knowledge speakers have of the sound patterns of their native language? To what extent, and in what manner, should phonetics be called upon to *explain* empirical generalizations about the typology of sound patterns and phonological systems? What bearing does all of this have on the question of whether there exists a phonological module of Universal Grammar (UG) and what elements of innate knowledge such a module might consist of?

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

The notion that the typology of phonological well-formedness is shaped by phonetic content, mediated by some substantive notion of "markedness," was an uneasy fit in classical models of generative phonology (cf. the famous Chapter 9 of Chomsky and Halle 1968). In the 1980s and early 1990s, phonetic content became more closely reflected through elaborate three-dimensional representations partly mirroring the structure of the vocal tract (feature geometry). Attempts were made to encode markedness in terms of formal complexity at the level of segments (radical underspecification, privative features) and of phonological rules (simple spreading and delinking operations vs. more substantial transformations); see Kenstowicz (1994a) for a review. These developments went hand in hand with the appearance of explicit generative models of the phonetics-phonology interface (Keating 1990), an increased interest among phonologists in interface issues (Archangeli and Pulleyblank 1994), and the emergence of "laboratory phonology" as a research field in its own right with a dedicated biannual conference series from 1987 onwards (Kingston and Beckman 1990). With the advent of Optimality Theory (Prince and Smolensky 2004; see McCarthy 2007d), well-formedness constraints on surface representations have become the "prime movers" in defining and circumscribing the sound patterns of languages, and the study of typological variation – modeled as permutations of a rank-ordered list of constraints – has taken on a far more central role in phonological theory. This has opened up the possibility of allowing synchronic grammars to incorporate functional explanations in terms of phonetic factors (cue availability, perceptual distinctness, articulatory effort) in a far more direct and explicit way than earlier frameworks permitted (see, e.g. Steriade 1999b, 2001; Hayes 1999; Hayes and Steriade 2004, and many contributions in Hayes, Kirchner, and Steriade 2004; cf. also Boersma 1998).

The explicit functional orientation of a great deal (though by no means all) of the work carried out under the Optimality Theory rubric in the last decade or more has reawakened an old line of criticism (Anderson 1981; cf. Sampson 1975), namely that grammar-internal explanations are frequently redundant in that they ignore the availability of an alternative explanation – often independently motivated, and functionally well understood – situated in the diachronic dimension of language change. According to this view, if we know (or can infer) how a particular synchronic pattern came into existence as the end product of familiar kinds of diachronic events (e.g. sound changes), and if we have a sound understanding of the nature and causes of such events in the "real-world" conditions under which speech is produced and perceived by language users and language learners, then this is all that is needed. Consequently, we should not attempt to invent a parallel synchronic explanation for the same set of facts, couched in grammar-internal terms, especially if this requires positing new theoretical constructs which are then attributed to the innate endowment of humans by way of Universal Grammar. In short, functional-phonetic explanations are to be sought and defined in the diachronic domain, not attributed to the inherent "design properties" of the synchronic grammars internalized by speakers/learners. The position that *diachronic explanations* for phonological patterns are not only a powerful tool, but a preferable way of deriving both the characteristic phonetic

"naturalness" of phonology and the occasional "unnaturalness" of certain sound patterns, is the main focus of this review.

The claim that explanations for synchronic sound patterns should be sought in their diachronic histories is by no means novel (Baudouin de Courtenay 1895; see Anderson 1985). This idea has never gone away, though it has enjoyed varying degrees of popularity in the century or so that has passed since phonology began to emerge as a discipline in its own right. The review presented here does not aspire to historiographic completeness. In particular, I will largely ignore works in the comparative-historical linguistic tradition (where the validity of appealing to the past to explain the present has generally been taken for granted). Rather, the goal is to introduce the reader to those explanatory models of a diachronic-evolutionary bent which have figured most prominently in the scholarly literature on phonology and phonetics – in particular the listener-based sound change model of John Ohala and the more recent Evolutionary Phonology model of Juliette Blevins – and to highlight and summarize in broad strokes the ongoing debate about the potential role that diachronic explanation has to play in phonology.

2 Phonetics in Phonology: Synchrony or Diachrony?

As a simple illustration, let us consider a concrete example of a common sound pattern with a plausibly phonetic explanation: the neutralization of lexical stop voicing contrasts like /t/ : /d/ in preconsonantal position, whereby /d/ becomes devoiced to [t] (/. . . VdC . . ./ → [. . . VtC . . .], neutralizing with /. . . VtC . . ./). This has an obviously plausible phonetic explanation in terms of speech perception (Steriade 1999b). To the extent that stops in this position are not released into a following vowel, there are few and relatively weak perceptual cues available in this position on the basis of which a given alveolar stop can be reliably identified correctly as either [t] or [d], as compared to other positions where the /t/ vs. /d/ contrast is not neutralized (e.g. where a stop is released into a following vowel, /. . . VdV . . ./ vs. /. . . VtV . . ./). If we wish to explain why this type of sound pattern is so widespread in the world's languages, why voicing neutralization should preferentially target this kind of position, and why no languages appear to selectively neutralize voicing contrasts in the complement set of environments (e.g. in prevocalic position), several options are available to us. First of all, we might contend that the apparent correlation with perceptual salience is an irrelevant curiosity, and that such asymmetric "licensing" of phonological contrasts is defined in abstract-structural terms which are themselves entirely substance-free (for an approach along these lines to sound patterns involving nasality, see Ploch 1999). On this view, the observed asymmetries and typological gaps are a consequence of the inherent properties of an innate Universal Grammar (UG): languages with selective voicing neutralization in prevocalic position are simply synchronically impossible systems.

A second alternative is to accept the notion that the phonetic-perceptual factors are what *explains* the occurrence and prevalence of phonological patterns of this

type, and that this phonetic basis of phonological patterns and processes is itself an intrinsic aspect of the synchronic grammars of languages. Though this position in no way presupposes an optimality-theoretic conception of grammar (cf. Stampe 1972; Archangeli and Pulleyblank 1994), we may express it in an optimality-theoretic context as the assumption that the well-formedness constraints which encode synchronic-phonological knowledge are themselves motivated (and perhaps even formulated) in explicitly phonetic terms. There are two versions of this view. One is that the constraints in question are part of an innate constraint inventory specified by UG, and that the set of constraints contained in UG is "grounded" in phonetics (Kager 1999: 11; cf. Archangeli and Pulleyblank 1994). Another possibility is that constraints are not innate but constructed by the language learner, who draws upon a general and grammar-external body of "phonetic knowledge" (Kingston and Diehl 1994) that informs and restricts the space of constraints which learners can posit (Hayes 1999; Hayes and Steriade 2004; see also Smith 2004). Returning to our voicing neutralization example, the fact that such neutralization often selectively targets preconsonantal position, but never prevocalic position, is a direct consequence of the constraints that make up synchronic grammars. Grammars may thus contain constraints that specifically penalize contrastive voicing in hard-to-perceive positions (or constraints which require contrast preservation in easy-to-perceive positions, if obstruent voicing as such is what is penalized), either as a reflection of UG or through induction from general phonetic knowledge by language learners. Grammars with constraints which express preferences diametrically opposed to these are impossible; such constraints are not innately supplied in UG, nor do they conform with learners' implicit knowledge of phonetics.

The third alternative approach – the focus of this review – is to acknowledge the functional bases of phonological patterning in the phonetic factors of speech production and perception, but to place the locus of phonetic influence squarely in the diachronic dimension of language change, rather than in the synchronic-universal domain of grammars and their "design." Viewed from this perspective, recurrent sound patterns are the product of recurrent diachronic events (sound changes), which have their ultimate causes in the physical conditions under which speaker-listener interactions take place in language use and in language transmission across generations. On this view, voicing is neutralized in preconsonantal (as opposed to prevocalic) position not because some constraint to this effect is part of the innate endowment of humans, nor because learners are predisposed to posit only such constraints as are grounded in phonetics. Rather, languages will show some tendency to acquire such neutralization patterns for the simple reason that, in positions where distinctive voicing is hard for listeners (including learners) to detect, listeners/learners will be liable not to detect it, erroneously interpreting a preconsonantal voiced obstruent as being voiceless and encoding it as such in their mental representation of the word form in question. If and when the pattern caused by such recurring misinterpretations becomes entrenched, the result is a language with systematic voicing neutralization precisely in those positions where such neutralization is phonetically motivated. Languages with the opposite, phonetically unmotivated pattern (neutralization in prevocalic position)

will be impossible – or at least infrequent to the point of being unattested – to the extent that this pattern cannot be attributed to plausible diachronic trajectories in terms of sequences of known types of events (sound change, analogical change).

3 The Listener-Based Model of Sound Change

One of the most influential conceptualizations of how phonetic factors influence sound patterns through diachronic change is due to John Ohala and his listener-based theory of sound change (Ohala 1981, 1989, 1993, and numerous other works). According to this theory, sound change originates in the kind of event when a listener misperceives or misparses the acoustic signal produced by the speaker, arriving at a representation which differs in some respect from that intended and encoded by the speaker. The *phonologization* (Hyman 1976) of such misapprehensions on the listener's part thus provides a channel through which articulatory, aerodynamic, and acoustic-perceptual factors come to shape phonological systems. While Ohala is rarely explicit on this point, misperception would be most likely to take place in words with which the listener is relatively unfamiliar, and phonologization through misperception is therefore perhaps most likely to occur during language acquisition in childhood, though to a certain extent lexical acquisition continues throughout the lifespan of the adult language user as well (as do changes in pronunciation; cf. Harrington, Palethorpe, and Watson 2000, and see Section 5 below).

There is little doubt that something along these lines is what accounts for A > B sound changes in cases where A and B are acoustically similar but articulatorily discontinuous. These would include such context-free (unconditioned) shifts in place of articulation as /θ/ > /f/, /f/ > /x/, or /k^w/ > /p/, as well as the intriguing phenomenon of "rhinoglottophilia" whereby aspiration or breathy voice shifts to nasalization or vice versa (Matisoff 1975; Ohala 1975). More importantly, however, Ohala's theory also extends to context-sensitive (conditioned) sound changes, and in general to a wide array of changes which have traditionally been attributed either to aimless "articulatory drift" (Paul 1880) or to the complementary, and often conflicting, teleological goals of minimization of (articulatory) effort and maximization of (perceptual) clarity (Grammont 1933). It is generally recognized that in speech perception, the listener is constantly engaged in normalization, correcting for predictable variation in the acoustic signal in order to arrive at the representation intended by the speaker. To mention but one example, Beddor, Krakow, and Goldstein (1986) demonstrate that whereas nasalization has a lowering effect on the perceived quality of a vowel, listeners are able to factor out this lowering effect when there is a plausible contextual source present in the form of an immediately adjacent nasal consonant (for other examples of such coarticulatory compensation, see Beddor and Krakow 1999; Beddor, Harnsberger, and Lindemann 2002). Ohala's fundamental insight is to attribute a variety of sound changes to such normalization gone wrong, as it were, where a listener either fails to correct

for a contextual effect or wrongly attributes some intrinsic property of a segment to contextual influence.

Let us consider in some detail an example where the underlying phonetic factors are arguably aerodynamic (Ohala 1983b). When a voiceless stop is released into a following vowel or sonorant, the Voice Onset Time (VOT) typically depends on the degree of constriction of the following segment, and so does the amount of frication noise in the release burst itself. Thus, for example, we will normally find noisier bursts and longer VOT values in sequences like [tɹ] or [ti] than in, say, [te] or [ta]. The aerodynamic explanation for this state of affairs rests on the fact that in order for the vocal folds to vibrate (as in voicing), air must be able to pass through the glottis with sufficient velocity, and this is in turn dependent on there being a large enough drop in air pressure from the subglottal cavity to the oral cavity. During the closure phase of a stop like [t], oral pressure builds up until it equals subglottal pressure, but when the closure is released and air escapes, oral pressure falls abruptly. How rapidly it falls, and how long before the subglottal-oral pressure differential has reached the critical threshold value at which voicing becomes possible, will depend on the width of the channel through which air has to escape out of the mouth. Consequently, the narrower the constriction into which the stop is released, the longer it will take for enough air to have left the oral cavity so that voicing can commence; furthermore, the narrower the channel, the more turbulence will be generated as air flows through it. Hence a [t] will display a longer VOT and a noisier release burst when it is released into an approximant, glide, or high vowel than when released into a mid or low vowel. This predictable effect can make (plain, unaspirated) stops sound somewhat like aspirated stops or affricates in this environment. When a listener fails to blame this predictable phonetic effect on the context, and instead perceives it as an intrinsic property of the stop (as either aspiration or affrication), a sound change has occurred (technically, a "mini" sound change; see below for clarification). This, then, is what accounts for common sound patterns whereby stops are obligatorily aspirated or affricated before high vowels and glides (e.g. in Japanese, where /tɯ/ → [tsɯ], /ti/ → [tʃi], /tjV/ → [tʃV]).

As mentioned above, there are two ways in which listeners may be mistaken in their decoding of the acoustic signal: by failing to correct for a contextual effect, or by over-correcting, attributing an intrinsic property to contextual influence. This translates into a two-way typology of listener-induced sound changes. In *hypocorrective* change, a contextual effect is misinterpreted as an intrinsic property of the segment, segment sequence, or word in question. The most obvious instances of this process are assimilatory sound changes, for example in vowel harmony (as well as umlaut, metaphony, etc.), which are the phonologized reflection of earlier V-to-V coarticulation patterns (Ohala 1994a, 1994b; see Przezdziecki 2005). Numerous other phenomena also fall under the hypocorrection rubric, such as the development of aspiration and/or affrication before high vowels and glides as described above. Tonogenesis (the emergence of distinctive tones) originates in the well-known phonetic effect of contextual pitch perturbations conditioned by the laryngeal properties (e.g. voicing) of an adjacent consonant (Hombert,

Ohala, and Ewan 1979). The emergence of "excrescent" stops in cases like English *prin*[t]*ce* (cf. also *Thompson* < *Thom+son*) is due to coarticulatory overlap of the oral closure gesture (for the nasal) and the velic closure gesture (for the following [s]), which creates the percept of an intervening oral stop homorganic with the nasal (Ohala 1981). Similarly, articulatory overlap can mask the percept of the middle consonant in a CCC sequence, such that a phrase like *perfe*/kt m/*emory* can become misperceived as *perfe*[k m]*emory*, even when the "deleted" /t/ is in fact being fully articulated by the speaker (Browman and Goldstein 1990) – a likely explanation for the high incidence of consonant deletion in such environments. Compensatory lengthening can likewise be attributed to hypocorrective sound change through the phonologization of pre-existing subphonemic differences in vowel duration (and/or in the duration of V-C transition cues), the phonetic sources of which may be quite varied (Kavitskaya 2002).

The other half of Ohala's sound change typology is *hypercorrective* change. In situations where the context provides a plausible source for some phonetic property of a segment's realization, the listener may erroneously "undo" this aspect of the segment. The most obvious instance of hypercorrective change is dissimilation. For example, when a phonemically labialized consonant is adjacent to a rounded vowel, a listener may attribute the labialization on the consonant to coarticulatory influence from the adjacent vowel, and interpret it as a (phonologically) non-labialized consonant. A well-known case is Classical Greek (/lukos/ < */luk^{w}os/ 'wolf', /kuklos/ < */k^{w}uk^{w}los/ 'wheel'); the deletion of postconsonantal /w/ before rounded vowels in English has a similar explanation (*sword* /soɹd/ < /swoɹd/, *two* /tu/ < /twoː/). A notable aspect of this theory of dissimilation and other hypercorrective sound changes is that they are predicted to be structure-preserving (Ohala 1993; Kiparsky 1995; though see Blevins and Garrett 1998: 519–520 for arguments against this position). In order for it to be possible for a listener to (mis)interpret an instance of [A] as being [B]+coarticulation, a [B] must already exist independently in the system.[1]

Two points about Ohala's listener-based theory are worth making here. Firstly, it should be emphasized that Ohala's explanations pertain, strictly speaking, only to the *initiation* (actuation) phase of sound changes. The misperception events on which the theory rests, and which have been successfully replicated in the laboratory in a variety of experiments, take place in a particular listener's perception of a particular utterance during some particular communicative event. The theory has nothing specific to say about how such a "mini sound change" (Ohala 1993) takes hold and spreads throughout a community. Nor does it account for how the effects of this "mini sound change" come to be manifested across the entire lexicon rather than just in the one word in question. With respect to the latter question, proponents of Ohala's theory typically appeal to some version of *lexical diffusion*, whereby sound changes are assumed to spread gradually through the lexicon, one word at a time (Wang 1969; Phillips 1984; Labov 1994). On this view, the exceptionlessness so characteristic of sound changes (as per the well-known "Neogrammarian principle") merely represents the endpoint of a gradual process of lexical diffusion (Krishnamurti 1998).

Second, the explanation of a sound change A > B in terms of perceptual error, where [A] (in some context) is misperceived as [B], does not imply that sound changes of the reverse type (B > A) ought to be equally possible or equally frequent. In some cases, such symmetry is expected; /k^{w}u/ > /ku/ (through hypercorrective misperception) and /ku/ > /k^{w}u/ (through hypocorrective misperception) are both possible, and both types of sound change are indeed attested. However, as has been amply revealed in perception studies, perceptual confusion is very often asymmetric. For example, Guion (1998) found that when listening conditions were artificially degraded so as to increase confusion rates, misidentifications of /ki/ as /tʃi/ occurred three to four times as often as the converse confusions (/tʃi/ heard as /ki/), which were quite rare (see also Winitz, Scheib, and Reeds. 1972; Plauché, Delogu, and Ohala 1997).[2] Consequently, the bizarre and unattested change /tʃi/ > /ki/ is predicted to be impossible, or at least extremely unlikely. The unsupported assumption that misperception necessarily be symmetric is occasionally encountered in works critical of phonologization accounts in terms of listener-based sound change (e.g. Steriade 2001).

The greatest success of the research program initiated by John Ohala and continued by many of his students and associates has been the identification of articulatory, acoustic and/or aerodynamic factors underlying a great number of common and less common synchronic sound patterns. In particular, these results have led to an increased awareness on phonologists' part of the fundamental role that speech perception has to play in explaining phonological patterns and processes (see, for example, many of the contributions in Hume and Johnson 2001b and Hayes *et al.* 2004).

4 What is the Role of the Speaker in Sound Change?

In sharp contrast to the traditional view in historical linguistics, Ohala's model assigns the speaker at best a tangential role in initiating sound changes and in influencing the direction such changes take. In the strongest version of the listener-based theory, the speaker's only role is to contribute to the "pool of synchronic variation" (Ohala 1989), from which listeners draw their conclusions about what mental representation underlies the speech signal. There is reason to believe that in its purest form this view is overly simplistic, and that speech production factors play a somewhat more direct role in shaping sound change – and hence, by transitivity, the typology of synchronic sound patterns that result from such change.

For example, the typology of recurrent sound changes is not always consistent with perception studies in the way that one would expect if all phonetically based sound patterns were ultimately due to misperception. Consider, for example, the extremely common process of nasal place assimilation, whereby a nasal takes on the place of articulation of an immediately following consonant (/m+d/ → [nd], /n+k/ → [ŋk]). This is consistent with the relative weakness of acoustic cues for place of articulation in this environment, particularly in nasals (Ohala 1990a). Indeed, Hura, Lindblom, and Diehl (1992) found that in heterorganic VC$_1$#C$_2$V

sequences, where C_1 ranged between nasal, fricative or stop, the place of articulation of C_1 was misidentified far more often than that of C_2, and most often when C_1 was a nasal. Misperception is thus a plausible explanation for the regressive directionality of assimilation, and for the particular susceptibility of nasals to such assimilation. However, the *assimilatory* character of the process is itself not supported by the experimental findings: when subjects misheard the place of articulation of C_1, they typically perceived it as consistently alveolar, regardless of C_2 place. Hura *et al.*'s (1992) alternative interpretation of assimilation in such clusters is that it constitutes "perceptually tolerable articulatory simplification" on the part of the speaker, not misperception of a heterorganic cluster as homorganic on the part of the listener. This interpretation is set against the backdrop of Lindblom's "H & H theory" of variation in speech production (Lindblom 1990), in which speakers are tacitly aware of factors which may affect listeners' ability to correctly perceive the signal, and actively tune their own production to compensate for those factors (Lindblom *et al.* 1995; see also Steriade 2001). On this view, then, the actual change from an unassimilated cluster to an assimilated one is initiated by the speaker, not the listener (though see Section 5 below for a modified version of this hypothesis).

This highlights another important aspect of Ohala's listener-based theory of sound change, namely its fundamentally *non-teleological* character (see Blevins and Garrett 2004 for recent discussion). In Ohala's model (unlike that of Lindblom), sound change does not serve any "purpose" or "goal," such as to make a word simpler or less effortful for the speaker to produce or easier for listeners to perceive reliably. Consequently, neither do the resulting synchronic sound patterns reflect "optimization" on such functional parameters in any meaningful sense of that term. Though teleological explanations have traditionally been eyed with great suspicion in historical linguistics (*pace* Grammont 1933; Martinet 1955; see, e.g. McMahon 1994; Lass 1997), the output-oriented and constraint-prioritizing character of Optimality Theory means that phonological derivation – and hence the application of phonological processes – is construed as an inherently goal-oriented procedure. Consequently, the vehemently anti-teleological stance inherent in Ohala's phonologization approach to the phonetic explanation of sound patterns has created a deep conceptual divide between works advocating this approach and ones closer to the current mainstream of generative phonological theory. A good example is the treatment of various positional neutralization phenomena in Barnes (2006) vis-à-vis the Optimality Theory analyses developed by Crosswhite (2001) and Smith (2005).

Though much work remains to be done in this area, various experimental studies appear to support the notion of listener accommodation in speech production. For example, the pronunciation of a word is affected not only by factors like frequency and relative information content (contextual (un)predictability; Lieberman 1963), but also by lexical neighborhood density. Words from high-density neighborhoods in the mental lexicon, which are thus potentially "hard" for the listener from the point of view of word recognition and lexical retrieval, appear to show hyperarticulation effects in production (Munson and Solomon 2004; Wright 2004)

as well as greater amounts of coarticulation (Scarborough 2004). Albeit suggestive, such findings are not always an unambiguous indication that speakers actively strive to make the listener's task easier. Pierrehumbert (2002) offers a conjectural alternative account of Wright's (2004) findings of vowel hyperarticulation in high-density words, which is couched in an exemplar-based model (on which, see Section 6 below). Pierrehumbert's suggestion is that it is, instead, listeners who are being selective in preferentially encoding hyperarticulated tokens of such words in their episodic memory; this selective storage in turn ends up biasing their own future productions of the words in question.

In the usage-based model advocated by Bybee (2001, 2007), the speaker is afforded a considerably more immediate and active role in the initiation of sound change than in Ohala's listener-based model. Bybee supports Mowrey and Pagliuca's (1995) claim that most sound changes involve articulatory *reduction* of one kind or another (in fact, Mowrey and Pagliuca hold this to be categorically true of all sound changes). Such reduction may be *substantive*, in that the magnitude of one or more articulatory gestures is reduced, or it may be *temporal*, in that sequences of articulatory gestures are compressed or overlapped in time. (Note that it is possible for a particular sound change to involve a combination of both types of reduction at once.) This corresponds rather closely to the kinds of gestural reduction and overlap effects that have been documented for casual speech (Browman and Goldstein 1990). The motivation for these sorts of articulatory reductions to occur over time is assumed to be the very general tendency for repeated and highly practiced motor behavior to result in reduction and temporal compression: "[w]ith repetition, neuromotor routines become more compressed and more reduced" (Bybee 2001: 78). In this conception of the mechanism by which sound change originates, the listener is clearly sidelined in comparison to the speaker. At most, the listener influences (passively and indirectly) the extent to which compression and reduction will still allow successful communication to occur (recall the Lindblomian notion of "perceptually tolerable articulatory simplification," discussed earlier).

Although Mowrey and Pagliuca (1995) go to great lengths to argue that many apparent examples of acoustically/perceptually motivated sound changes can be reinterpreted in terms of temporal or substantive reduction, there are certain types of change which are hardly amenable to this kind of explanation, and for which Ohala's listener-based account is almost certainly correct. In particular, this is true of changes which involve abrupt articulatory discontinuities (see Section 3 above), such as when labiodental [f] turns into velar [x] (or vice versa), or when breathiness or aspiration turns into nasalization. The same applies to phenomena like tonogenesis, in which low-level and largely mechanical pitch perturbations spawn distinctive phonological specifications for tone on a neighboring vowel. It is difficult to see how adopting the notions of temporal and substantive articulatory reduction would help us better understand how and why such changes occur.

Bybee (2001) is skeptical of Ohala's general account of sound change as misperception, arguing that "it is unlikely that hearers who have already acquired the phonetics of their dialect would misperceive already acquired words to the extent

that that might cause a sound change" (p. 75). She does, however, concede that certain sound changes might originate in children's misperception or misanalysis of the acoustic signal at the acquisition stage. Most interestingly, though, Bybee argues (following Phillips 1984, 2001) that articulatorily and perceptually motivated sound changes are expected to interact very differently with word frequency, and should therefore leave distinct lexical diffusion "footprints." Given that the compression and reduction inherent in (speaker-based) articulatorily motivated changes is due to the articulatory automation of often-repeated elements, it would follow that such changes should affect *high*-frequency forms first, and only later spread to forms with lower token frequency. By contrast, perceptually motivated (listener-based) sound changes, which are essentially a form of imperfect learning, ought to affect *low*-frequency forms first, since the listener/learner has much greater exposure to more frequent forms and is therefore likely to perceive and acquire those correctly. It remains to be seen whether this prediction stands up to scrutiny, and what its implications might be for analyses of synchronic systems and phonological typologies. In recent years, probabilistic aspects of sound patterns (e.g. variation and optionality, gradient well-formedness judgements, probabilistic phonotactic restrictions, word-frequency effects) have come to figure more and more prominently within mainstream phonological theory; see Coetzee and Pater (this volume) for a recent overview.

It should be noted that Bybee (2001: 75) also acknowledges that the outcome of prior (speaker-based) articulatory compression and reduction changes may in turn be subject to (listener-based) perceptual reanalysis later on. For example, a word-final vowel+nasal sequence might undergo considerable gestural overlap (resulting in nasalization of a substantial portion of the vowel) and articulatory reduction of the nasal (resulting in shorter duration, and perhaps less-than-complete oral closure). As a result of these articulatory changes, the misperception of the vowel+nasal sequence as a mere nasalized vowel becomes much more likely than otherwise. In practice, the difference between this account and Ohala's purely listener-based treatment is largely a matter of emphasis. For Ohala, the gestural reduction and overlap is merely present in the "pool of synchronic variation" (Ohala 1989) on which listener-based sound changes operate (i.e. these articulatory properties happen to be found in some subset of the production tokens that make up that "pool"); it is only the perceptual reanalysis event that is deemed to deserve the label "sound change." For Bybee, the focus is instead on the processes by which such articulatorily reduced production variants come into existence in the first place; that process is viewed as the more significant "sound change" event, rather than any perceptual reorganization that may ensue. This difference of perspective also relates to what appear to be somewhat different assumptions about the nature of phonological representations and lexical storage. To the extent that this can be inferred from Ohala's writings, he is for the most part presupposing something akin to a (quasi-)phonemic representation of word-forms (though probably with canonical, contextually predictable allophonic variants spelled out). Bybee's usage-based model, on the other hand, rests on a rich-storage conception of the mental lexicon, in which "each word is represented with a range of variation

that corresponds to the actual tokens of use experienced by the user" (2001: 42), such that "phonological representations [of lexical items] are based on categorized tokens of use" (2001: 62). We shall return to this issue in the following section in the context of Blevins' Evolutionary Phonology model, as well as in the discussion of exemplar-based models in Section 6 below.

To sum up, while much is yet unclear about the relative contributions of the speaker and the listener to diachronic sound change, one thing has become abundantly clear: the "pool of synchronic variation" in speech production is far from random, but rather intricately structured and influenced in principled ways by a variety of factors. Any explicit theory of sound change, and any theory which seeks to explain the properties and cross-linguistic typology of synchronic-phonological systems with reference to the diachronic domain of language change, needs to take into account this complex interplay of production and perception in speaker-listener interactions, both in fully competent adults and in children acquiring language.

5 Sound Change in Evolutionary Phonology

As part of her Evolutionary Phonology model, Blevins (2004; see also Blevins 2005, 2006c; Blevins and Garrett 1998, 2004) outlines a model of sound change which is essentially an elaboration of Ohala's earlier model, and which in part addresses some of the issues raised in the previous section. In this "amplified" model of listener-based sound change, referred to as the "CCC model," sound changes come in three basic varieties, dubbed CHANGE, CHANCE, and CHOICE. While Blevins does acknowledge Ohala's hypo- vs. hypercorrection dichotomy, that distinction plays no explicit role as such in her model.

In both of the first two categories of sound change, CHANGE and CHANCE, the listener (re)constructs a phonological representation which differs from that intended by the speaker, much as in Ohala's hypo- and hypercorrective misperception scenarios. However, it is only in CHANGE that some aspect of the acoustic signal is actually being misheard outright. For example, the listener might hear [v] when the speaker in fact produced [ð] (or, similarly, [kʲi] might be misheard as [tʃi], [kʷu] as [pu], or breathy-voiced [a̤] as nasalized [ã]). Consequently, CHANGE automatically and immediately leads to a change in pronunciation of the word in question by the listener (as compared to that of the speaker).

The sound changes Blevins labels as CHANCE, by contrast, are ones in which the acoustic signal, and the sequence of articulatory events it reflects, is itself genuinely ambiguous in some way, and where the listener simply happens to parse the signal in a way that diverges from the speaker's intended representation. For example, the speaker may produce /. . . kʷu . . ./ as [. . . kʷu . . .], with phonemic labialization on the velar stop, whereas the listener, correctly hearing this [. . . kʷu . . .], incorrectly parses it as /. . . ku . . ./, erroneously attributing the labialization on the [kʷ] to coarticulation with the following rounded vowel (exactly as in the discussion of hypercorrective change in Section 3 above). It is important

to note that this misinterpretation need not noticeably affect the listener's own pronunciation of the word, as the listener is likely to continue to render this /. . . ku . . ./ as [. . . k^{w}u . . .], at least under most circumstances; CHANCE is thus largely covert. In the vast majority of cases, CHANCE involves features or segments which have temporally elongated acoustic cues or which are otherwise difficult to localize within the segmental string. It is therefore assumed to be the main mechanism responsible for the emergence of various feature-spreading processes, as well as many instances of metathesis and dissimilation.

For example, in Tsilhqot'in, vowels undergo retraction and/or lowering next to both uvulars and pharyngealized sibilants, and such effects on vowel quality provide the only reliable perceptual cues to the pharyngealization contrast in sibilants (see Hansson 2007b and sources cited therein). In an intended sequence like /. . . $S^{\text{ˤ}}$VQ . . ./ (where $S^{\text{ˤ}}$ stands for some pharyngealized alveolar sibilant, and Q for some postvelar consonant), the retracted and lowered quality of the intervening vowel becomes ambiguous as to its source. The listener needs to decide whether to infer (correctly) that the vowel quality is partly due to the preceding sibilant – and hence that the sibilant in question must be a phonologically pharyngealized one – or instead to conclude (incorrectly) that the lowering/retraction is entirely due to the following uvular, and that the sibilant must thus be of the *non*-pharyngealized kind. In other words, should a pronunciation like [-ts'ɑχ] 'sinew' be interpreted as /-ts$^{\text{ˤ}}$'æχ/ or as /-ts'æχ/? As it turns out, all morphemes which were originally of the shape -$S^{\text{ˤ}}$VQ have been historically reanalyzed as -SVQ in Tsilhqot'in (Krauss 1975; see Hansson 2007b), in what amounts to a CHANCE-based dissimilatory sound change.

The third type of sound change, CHOICE, is where Blevins' model departs most clearly from Ohala's, and where the speaker becomes implicated to a certain extent. Blevins adopts certain aspects of the H & H theory of Lindblom (1990), viewing variation in production as falling largely on a continuum from relatively hypoarticulated variants to relatively hyperarticulated ones. A speaker thus produces, and a listener is exposed to, a range of multiple variant pronunciations of individual word forms. For example, the realizations of /. . . ut' . . ./ might range from the relatively "hyper" [. . . ut' . . .], via intermediate [. . . u̟ˀt . . . , . . . u̠ˀt . . . , . . . ʏˀt . . .], to the relatively "hypo" [. . . ʏʔ . . .], with shorter vowel duration and hence greater centralization/fronting and coarticulatory overlap, as well as near-complete reduction of the oral gestures of the /t'/. If the relative frequency distribution of these variants changes, listeners may come to choose a different phonological representation to represent this continuum.[3] For example, if variants with considerable vowel fronting become particularly frequent in the ambient production data ([. . . ʏˀt . . .], [. . . ʏʔ . . .]), the listener might choose /. . . yt' . . ./ as the lexical representation, or even /. . . yʔ . . ./ if variants with (largely) debuccalized renderings of /t'/ are particularly common.

Blevins (2004) does not cite Hura *et al.* (1992) and the problems posed by their finding that typical patterns of misperception in heterorganic consonant clusters fail to match the most typical sound change affecting such clusters, namely assimilation (see Section 3 above). Following Ohala (1990a), Blevins clearly considers

the primary mechanism for regressive place assimilation in C_1C_2 clusters to be CHANGE, that is, outright misperception, a view which would appear difficult to reconcile with Hura *et al.*'s findings. She does, however, note (pp. 116, 118–119) that another contributing factor may be "coarticulation," by which she means the (near-)complete masking of the oral gesture of C_1 through gestural overlap (Browman and Goldstein 1990), such that /. . . np . . ./ might be pronounced, roughly, as [. . . n͡mp . . .] with little or no audible trace of the alveolar closing gesture in the acoustic signal. Note that the transcription [n͡m] is here intended to convey a doubly-articulated nasal, in which the alveolar and bilabial closures are essentially cotemporaneous; in other words, [. . . n͡mp . . .] represents a /n/+/p/ cluster in which the labial closing gesture for the /p/ starts already at the onset of the nasal, and thus overlaps the alveolar closure of the /n/. Blevins (2004: 44) seems to have this in mind when she states that "CHOICE may also be involved" in some cases of assimilation.

To be precise, invoking CHOICE here would by definition have to mean that such gestural-overlap tokens, having become gradually more frequent in the ambient distribution of production variants to which listeners are exposed, become the basis for positing a phonological representation different from the original one (/. . . mp . . ./ rather than /. . . np . . ./). Nevertheless, it would seem that CHOICE alone is not sufficient even here, and that in order to get from the dual-gesture [. . . n͡mp . . .] to the consistently single-gesture [. . . mp . . .] = /. . . mp . . ./, CHANCE needs to be invoked as well (rather than CHANGE, since the acoustic signal is presumably ambiguous as to the [. . . n͡mp . . .] vs. [. . . mp . . .] articulatory distinction). In other words, the two alternative mechanisms are either CHOICE+CHANCE or else pure CHANGE (the latter being contradicted by the findings of Hura *et al*. 1992). A third alternative possibility would be to embrace more fully Lindblom's notion of the "hypo–hyper" continuum of production variants as being generated by deliberate "perceptually tolerable articulatory simplification" on speakers' part (see Section 4 above). It is at least conceivable that fully-assimilated realizations (i.e. [. . . mp . . .], not just "coarticulated" [. . . n͡mp . . .]) are innovated as "hypo" production variants by the *speaker*, in which case CHOICE alone would be entirely sufficient. If nothing else, this example demonstrates how difficult it can be to disentangle the various possible mechanisms of change, and the degree to which the listener and the speaker are each implicated in those mechanisms, when individual cases are considered in detail.

Finally, one consequence of including CHOICE in the model is that this explicitly allows for changes in the phonetic realization of words – for example, in the relative frequency of pronunciation variants – over the lifespan of the speaker (for an amusing example, see Harrington *et al.* 2000). In other words, it is not assumed that all sound change necessarily constitutes imperfect learning at the acquisition stage. This in turn opens the door to word-frequency effects on sound change such as those addressed by Bybee (2001), as discussed in Section 4 above. In this respect, Blevins' Evolutionary Phonology framework is closely aligned with exemplar-based models of speech production and perception, to which we now turn.

6 Exemplar-based Models and Simulations of Sound Pattern Evolution

Traditional generative models of the phonetics-phonology interface (e.g. Keating 1990; Coleman 1998b) take the view that phonological representations (underlying and surface representations alike) are composed of discrete symbolic elements – possibly redundancy-free as per some version of Underspecification Theory – which are mapped or "transduced" onto the continuous/analog/gradient domains of articulation and acoustics. In recent years, this traditional view has increasingly been challenged by probabilistic exemplar-based models of speech perception and production (Johnson 1997; Pierrehumbert 2001a, 2002; see also Bybee 2001). In such models, lexical entries are represented by clouds of exemplars in episodic memory. An exemplar is essentially a memory trace representing an individual token previously encountered in perception (exemplar models are also known as multiple-trace models; Hintzman 1986); some decay function is assumed, such that older tokens fade over time. In perception, stored exemplars are activated probabilistically in proportion to their similarity to the token under consideration, which is then categorized in accordance with these. In production, a subset of exemplars is selected and a production motor plan is arrived at by averaging over these selected exemplars (in proportion to their activation levels).

The production-perception feedback loop inherent in exemplar-based models entails that exemplar clouds – of entire word forms, as well as of the multitude of categories that cross-classify such whole-word or whole-utterance representations – can and will be subject to gradual change over time. Language is thus viewed as a fundamentally dynamic and usage-based system (Bybee 2001, 2007; Silverman 2006a), and the emergence and evolution of sound patterns at various levels of granularity can be analyzed in terms of formal and mathematically explicit models of these complex dynamics in acquisition, in the competence of the adult speaker, and in speaker-listener interactions. For example, Wedel (2004, 2006) uses agent-based computational simulations to demonstrate how the merger, maintenance and/or transformation of phonological contrasts over time can be derived in exemplar-based terms (see also Pierrehumbert 2001a, 2002). An instantiation of the same general approach is outlined by Silverman (2006a), who also shows how an evolutionary exemplar model can explain the preservation and even exaggeration of contrasts over time. Silverman (2006a: 135–143, 2006b) discusses a particular sound change in Trique, by which lip rounding has spread rightwards across velar consonants ([. . . uga . . .] > [. . . ug^{w}a . . .]). He argues that this development is fundamentally a matter of *contrast enhancement*: what was once an [uga] : [uda] contrast has been replaced with a more acoustically distinct [ug^{w}a] : [uda] contrast. (As no [ug^{w}] or [uk^{w}] sequences existed in the language prior to this development, the increased category separation did not incur any extra "cost" in terms of other contrasts being infringed upon.) Importantly, Silverman rejects any kind of teleological, speaker-oriented interpretation of the mechanism by which this change occurred (cf. the discussion in Section 4 above).

That is, at no point were any speakers actually striving to make their productions more distinctive or easier to parse. Rather, Silverman suggests a listener-based account "by which contrasts might be enhanced passively, evolving over generations of speakers, due to the communicative success of some tokens, and the communicative failure of others" (Silverman 2006b: 141).

The core of Silverman's proposal is that those tokens of /uga/ which have extensive coarticulatory rounding ([ug^{w}a]) are more likely to be successfully perceived and categorized correctly than ones with little or no coarticulation ([uga]); a certain percentage of the latter will be misperceived/miscategorized, and will therefore fail to be added to the exemplar cloud representing /uga/. This very small but persistent bias will continue to assert itself over cycles of speaker-listener interactions and across generations of learners. Over time, [ug^{w}a] exemplars will thus gradually gain ground at the expense of [uga] tokens, until they come to dominate the exemplar cloud, at which point they effectively constitute the norm. By contrast, tokens of /uda/ with coarticulatory rounding on the intervening consonant ([ud^{w}a]) fare worse than ones with less rounding ([uda]), as they are *less* likely to be perceived and categorized correctly (i.e. more prone to being misheard as [uga]); therefore, the same mechanism will have the effect of inhibiting excessive coarticulatory rounding in non-velar consonants.[4]

In response to what she views as problems for simple exemplar-based models of lexical representation, Bybee (2001: 138–143) outlines what she calls a "modified" exemplar model. She notes that contextual variants of words are typically not very stable, and that reorganization of variant distributions may occur. For example, in dialects of Spanish where [s] alternates (variably) with [h] and ∅ as realizations of /s/ in preconsonantal (and, to a more limited extent, prepausal) environments, one may find (e.g. in Cuban Spanish) that in words with final /s/, the frequency of variants with [h] or ∅ has increased dramatically in __#V contexts, under the influence of __#C contexts (in which [h] and ∅ variants dominate), while still remaining quite rare in prepausal contexts. The influence exerted by the __#C pattern is clearly related to the fact that the vast majority of Spanish words begin in a consonant; an /s/-final word will therefore find itself in a __#C environment more than twice as often as it will occur in a __#V environment. Bybee notes that if exemplars are stored with their contexts specified, the occurrence of (a particular type of) exemplar in the "wrong" environment is unexpected, and argues (2001: 142) that "[t]he stabilization of a single variant for a word suggests that representations are exemplar-based, but that all exemplars are not equally accessible." Her (somewhat sketchy) proposal is that (sub)sets of similar exemplars – regardless of the contexts from which these may be drawn – are mapped onto a single representation. This effectively maps the exemplar cloud to a relatively small set of exemplar "types," which are in principle context-free, and some of which are more frequent overall than others. These "more central" exemplar types "are more accessible and may replace the more marginal ones" over time (Bybee 2001: 142). In the Cuban Spanish example above, [h] variants of words with final /s/ were more central/frequent overall than [s] variants; as a result, the former have ended up infiltrating the __#V contexts in which [s] variants had been predominant.[5]

Though the question of such "variant reorganization" certainly deserves attention and explicit analysis, it is not so obvious that the sorts of developments that Bybee discusses cannot be accommodated in standard exemplar models. The issue would appear to hinge on the question of exactly how individual exemplars are cross-classified and categorized along a number of intersecting parameters, and how that cross-classification is accessed and made use of in production. Exemplar-based production models like that of Pierrehumbert (2001a, 2002) typically assume that in the production of a given category in a given environment, stored exemplars of that category from *any* environment can in principle be activated and contribute to the calculation of a production target (with exemplars from more similar environments perhaps being more strongly activated than others). For this reason, it seems that a mechanism is already in place by which the more numerous __#C exemplars (among which [h] predominates) can gradually come to influence the incidence of [h] variants in __#V contexts as well.

In fact, this is exactly how exemplar-based models account for the otherwise puzzling phenomenon of *incomplete neutralization* (Port and O'Dell 1985; Warner, Jongman, Sereno, and Kemps 2004) and the related issue of near-mergers (Labov, Karen, and Miller 1991; Yu 2007b). For example, in their study of neutral vowels in Hungarian vowel harmony, Benus and Gafos (2007) demonstrate how the phonetic realization of neutral [i, iː, eː] in those monosyllabic roots which idiosyncratically take back-vowel suffixes ([iːr] 'write', cf. [iːr-n̺k] with the dative suffix) is slightly but significantly more back than that of the same vowels in analogous front-harmonic roots ([hiːr] 'rumor', dative [hiːr-nɛ̺k]). This is true even in isolation, when no suffix follows the roots in question. The explanation for this subtle and subphonemic "contrast" is that in the production of an unsuffixed form like [iːr], exemplars from suffixation contexts (e.g. [iːr-nɔk]) are activated as well. In those latter tokens, the [iː] is subject to coarticulatory backing due to the nearby back vowel, and therefore has a more retracted realization. Under the influence of such retracted-[iː] exemplars, the average production of a back-harmonic root like [iːr] in isolation gets shifted towards a slightly more back version of [iː] (as compared to that of a front-harmonic root like [hiːr]). The same explanation applies to the incomplete neutralization frequently observed in word-final devoicing processes in languages like German or Dutch (Port and O'Dell 1985; Warner *et al.* 2004). The phonetic realization of the word-final stop in German *Rad* "wheel" (traditionally transcribed [ʀaːt]) is influenced by the voiced realization in exemplars of that same word from suffixed contexts (e.g. genitive *Rad-es* [ʀaːdəs]), such that this devoiced "[t]" is in fact not quite phonetically identical to the genuinely voiceless [t] of *Rat* 'counsel'. If exemplar-based models are able to accommodate this sort of cross-context transfer, then they should have no trouble dealing with the kinds of diachronic changes that lead Bybee (2001) to propose what amounts to a hybrid between an exemplar model and a prototype model.

In exemplar-based models of the kind outlined above, known functional factors – exigencies of production, perception, and processing – can act as constant biases exerting pressure on the system over time (Pierrehumbert 2001a; Wedel 2007). The distribution of exemplars may shift under the influence of such factors, which

in turn provides the seeds for potential (larger-scale) sound changes of familiar kinds. Moreover, language-specific frequency distributions may also provide a source of biases of a similar kind, with interesting consequences. For example, the results of the computational simulations by Wedel (2007) turn out to replicate closely Gordon's (2002b) typological observation that whether a given language categorizes CVC syllables as phonologically heavy (bimoraic) or light (monomoraic) is strongly, though non-deterministically, correlated with the ratio of sonorants to obstruents in the inventory of syllable codas in the language in question. In other words, the more sonorous a language's "average" coda is, the more likely that language is to attribute weight to *all* codas, regardless of their sonority.

The full explanatory potential of exemplar models is most profitably investigated with the help of computational implementations, often in the form of agent-based simulations, which help explore how sound patterns, representations, and contrasts are expected to *evolve* over time in such models, and to what extent these predictions match what is known about attested diachronic changes and synchronic-phonological patterns and typologies. The most notable representative of this strand of research is recent work by Wedel (2004, 2006, 2007). Other works which exploit the self-organizing capabilities of complex adaptive systems through agent-based simulations of diachronic change (though not necessarily explicitly couched in exemplar-based terms) include de Boer (2001) on vowel inventories and Harrison, Dras, and Kapicioglu (2002) on vowel harmony; see also Boersma and Hamann (2008) on contrast dispersion, which relates to the work on contrast maintenance and enhancement already cited (Silverman 2006a, 2006b; Wedel 2004, 2006). In a recent study, Wedel (2007) shows how positive feedback – "analogical error," in the form of a perception bias favoring previously encountered exemplars similar to the current token – can give rise to regularity across the lexicon. Given such reinforcing feedback, the categorical and regular patterns so characteristic of phonological systems end up emerging as stable states. This line of inquiry, whereby higher-order organizational aspects of synchronic phonologies are themselves viewed as emergent structures arising from the cumulative interaction of other, more basic factors, is likely to yield more interesting results in the future.

7 Synchronic Universals and the Adequacy of Diachronic Explanation

In the ongoing debate about the viability of diachronic-evolutionary approaches, such as that advocated by Blevins (2004), the central question is to what extent such approaches are adequate for correctly predicting observed cross-linguistic typologies, in particular as regards typological gaps and apparent universals. Models like Evolutionary Phonology explicitly allow for "unnatural" histories: diachronic trajectories which involve the telescoping of sequences of independent sound changes, or analogical processes like reanalysis or rule inversion (Blevins 2005; Garrett and Blevins 2009). Such models are therefore fairly permissive in the range of synchronic sound patterns they predict to be possible, though patterns

which depend on highly specific and fortuitous sequences of events for their emergence are expected to be quite rare, possibly to the point of being (as yet) unattested.

In many ways, the ability to account not only for phonetically "natural" sound patterns but also more arbitrary ones (the "crazy rules" of Bach and Harms 1972; cf. Anderson 1981) is one of the main strengths of Evolutionary Phonology and related explanatory frameworks. Typologically aberrant outliers can be accommodated – rather than ignored or explained away by ad hoc maneuvers – without undermining the soundness of the strong cross-linguistic generalizations that such systems so blatantly violate. For instance, Barnes (2006) calls attention to the anomalous unstressed-vowel reduction pattern in Seediq, where unstressed /e/ is merged with /u/ (as is /o/), rather than with /i/. As it turns out, the vowel which now alternates between (stressed) front unrounded [e] and (unstressed) back rounded [u] goes back to central /ə/ at an earlier historical stage. At the time when the unstressed-vowel mergers took place as phonetically driven sound changes, the change in question was thus not the highly unexpected [e] > [u] but rather the less unusual [ə] > [u]. Similarly, Hyman (2001a) argues that Tswana shows evidence of a high-ranked constraint against voiced stops in nasal+stop clusters, the effect of which is diametrically opposite to the extremely common (and "natural") process of postnasal voicing in the world's languages.[6] If correct, this contradicts standard assumptions of markedness and typological variation in Optimality Theory (for example, see Chapter 2 of Kager 1999 and sources cited therein). Hyman goes on to demonstrate how this synchronic state of affairs is the end product of a particular sequence of diachronic changes which are themselves quite unremarkable.

The most serious criticism raised against diachronic-evolutionary models is not that they permit such phonetically unmotivated and typologically aberrant systems, but rather that they do so all too easily and too indiscriminately. The question is whether there are synchronic patterns which should in theory be diachronically accessible, but which are in fact categorically unattested. Kiparsky (2006, 2008b) takes just this position, arguing that a set of universal design principles of synchronic grammar – Universal Grammar (UG) in the standard generative sense – constrains the possible outcomes of diachronic change. De Lacy (2006) refers to this constraining influence of UG on cross-linguistic typologies as "straitjacket effects" (see also de Lacy and Kingston 2006). Both Kiparsky and de Lacy accept that the functional pressures of language use do to a certain extent determine typological generalizations, statistical tendencies, and frequencies, and acknowledge the validity of seeking such diachronic-functional explanations. However, they maintain that certain *true* universals also exist, which cannot be adequately accounted for in this manner, and which must therefore owe their existence to properties of the phonological component of UG. In other words, diachronic explanations alone are not sufficient.

A topic which has figured prominently in this discussion is the typology of final voicing neutralization, and whether such neutralization exclusively favors voiceless over voiced obstruents (as predicted by formal theories of markedness;

cf. Lombardi 2001). Lezgian has been claimed to show an active process of syllable-final voicing (Yu 2004a), and Blevins (2006c) adduces several additional cases, arguing that the cross-linguistic prevalence of devoicing over voicing in final neutralization is merely a (strong) statistical tendency dictated by recurrent patterns of sound change. Kiparsky (2006, 2008b) disputes Yu's and Blevins' interpretation of the phonetic and phonological facts of each of these cases, maintaining that the typological gap is in fact absolute and hence a true universal. Moreover, as Kiparsky (2006) points out, it is easy enough to imagine simple sequences of diachronic changes which ought to be capable of giving rise to synchronic patterns of final voicing rather than devoicing.

For example, consider a hypothetical language which originally had a geminate vs. singleton contrast in stops, such that /matt-/ and /mat-/ were distinct morphemes (e.g. realized as [matt-a] vs. [mat-a] when some suffix /-a/ is added). Let us further suppose that in this language, the length contrast was systematically neutralized by degemination in word-final position, such that unsuffixed /matt/ → [mat], surfacing identically to underlying /mat/ → [mat]. Now, if this language were to then undergo a lenition sound change whereby [t] > [d] and [tt] > [t] in all positions, and concomitant restructuring of the underlying contrast – such that the old /tt/ vs. /t/ opposition is now /t/ vs. /d/ instead – the result would be a system with word-final neutralization of the /t/ : /d/ contrast in favor of voiced stops. A root like /mad-/ (reflecting earlier /mat-/) would retain voiced [d] in all positions (suffixed [mad-a], unsuffixed [mad]). On the other hand, a root like /mat-/ (going back to earlier /matt-/) would display a [t] ~ [d] alternation: the stop would surface as voiced [d] in final position (unsuffixed [mad]) but as voiceless [t] elsewhere (suffixed [mat-a]).

Even if one accepts Kiparsky's argument, and his alternative interpretations of alleged instances of final voicing, it is not immediately obvious how one should envisage the intervention of UG in blocking the emergence of cases like the hypothetical one just described. One not so plausible interpretation is that UG has a "prophylactic" effect, quite literally blocking such a (phonetically driven) lenition sound change from ever taking place in a language with these properties. In an analogous language lacking final degemination, by contrast, the exact same sound change would presumably have been allowed to progress unhindered by UG. To the best of my knowledge, this is not a commonly held version of the view that synchronic typological universals exist independently of diachronic considerations.

A less extreme version of the view that synchronic universals constrain diachronic change is that the lenition sound change in question would still be able to take place, but that the resulting sound pattern (which would give the *appearance* of an active word-final voicing process) would either automatically trigger further changes or simply not be acquired by learners as an aspect of the synchronic-phonological grammar they construct. For example, the resulting distributional gap (absence of /t/ in final position) might thus be treated by learners as an "accidental" gap, waiting to be filled by loanwords and other new formations. Observed [t] ~ [d] alternations (as in [mat-] ~ [mad-]) would be lexicalized, dealt with by the learner in terms of listed allomorphs at the underlying level ({/mat-/,

/mad-/}), and should thus become highly susceptible to analogical leveling and decreased productivity over time.

Building on work by de Lacy (2002a, 2004) on universal markedness relations in sonority-driven stress systems, Kiparsky (2008b) argues that in such a stress system, a simple and seemingly innocuous sound change ought in principle – if synchronic properties of phonological systems were merely a product of their diachronic histories – to be able to give rise to a markedness reversal that subverts otherwise exceptionless typological universals. For example, consider a stress system like that of Gujarati (de Lacy 2002a), in which stress normally falls on the penultimate syllable ([apwána] 'to give', [ekóteɾ] '71'), but is attracted away to the antepenultimate or the final syllable if the vowel in that syllable is more sonorous than the one in the penultimate ([tádʒetəɾ] 'recently'), where sonority is defined in terms of the (partially conflated) hierarchy [a] > [e, o, i, u] > [ə] (with ">" standing for "is more sonorous than"). Kiparsky (2008b) points out that a simple context-free sound change [a] > [ə] (such as occurred in Sanskrit, for instance) would be capable of turning such a system on its head. As a result, [ə] might come to *attract* rather than repel stress, and would effectively end up counting as the most sonorous vowel for the purpose of stress assignment. However, no stress system with such properties exists, and in Kiparsky's view this is because it would constitute a synchronically impossible system. Importantly, he does not deny the possibility that an [a] > [ə] sound change could occur in a language like Gujarati, but rather claims that this would "destroy the phonological regularities of Gujarati's stress system, with the result that it would have to be reanalyzed with lexically marked stress" (Kiparsky 2008b: 51). In other words, learners confronting the resulting surface patterns of stress assignment would by necessity interpret stress as being unpredictable (or at least subject to a great number of lexical exceptions).[7]

Citing the same thought experiment on vowel change in a Gujarati-like sonority-driven stress system, de Lacy and Kingston (2006) take a much stronger position. Whereas Kiparsky (2008b) expects the sound change to necessitate *covert* reanalysis and lexicalization of the stress system on the part of language learners, de Lacy and Kingston go one step further and assert that it will – simultaneously and by blind necessity – trigger *overt* change in the stress patterns of the relevant words (i.e. ones where a stress-attracting [a] is turning into [ə]). They claim that in a language with sonority-driven stress, a sound change like [a] > [ə] "will necessarily alter the stress in words that have undergone the change" (de Lacy and Kingston 2006: 7–8). In other words, in a word like [tádʒetəɾ], "the *a > ə change must be simultaneous with the change in stress position to [tədʒétəɾ] – there is no stage in the language's history which would have [tə́dʒetəɾ]" (de Lacy and Kingston 2006: 8). It is difficult to see how this position could be integrated with current theories of sound change such as those of Ohala (see Section 3) or Blevins (see Section 5), not to mention exemplar models or usage-based models like that of Bybee (2001).

The more moderate interpretation, by which sound patterns that are synchronically "impossible" (i.e. violate synchronic universals) may arise but will simply

not be captured as such by learners, is not implausible in principle. For example, in the Northern dialect of Icelandic, aspirated and unaspirated voiceless stops contrast word-initially (/t^{h}au/ 'toe', /tau/ 'coma'), but intervocalically only the aspirated stops occur (/kaut^{h}a/ 'riddle'; Hansson 2003). The historical background is well understood: in postvocalic positions, the Modern Icelandic reflex of the Proto-Indo-European contrast from which the initial /t^{h}/ : /t/ distinction derives is instead /t^{h}/ : /ð/ (and similarly for obstruents at other places of articulation). Perhaps unsurprisingly, a steady stream of borrowings over the past few centuries has gradually been filling the distributional gap (/ratar/ 'radar', /lɛkou/ 'lego blocks'). This could be taken as evidence that the postvocalic "neutralization" of T^{h} : T in favor of aspirated stops was never an aspect of the Northern Icelandic phonological system in the first place, even at the earlier stage when the distribution facing language learners – that is, the (supposedly "accidental") gap – would have been completely exceptionless. Variant pronunciations in some items suggest that aspiration has occasionally been imposed, however (e.g. /t^{h}upa/ ~ /t^{h}up^{h}a/ 'tube'), which suggests that an account along these lines may be overly simplistic. In any case, the main problem with this view of UG–diachrony interaction is that it is extremely hard to falsify in principle. In practice it is all too easy to explain away apparent counterexamples (alternations or distributional patterns which violate some alleged synchronic universal) as being lexicalized, morphologized, or in some other way not belonging to the "real" phonology of the language.

From the point of view of a generative phonologist, true universals, if they exist, are restrictions on possible grammars: knowledge systems internalized by language learners. The assumption that there exists a generative phonological module of grammar brings with it certain complications, particularly if this module is assumed to manipulate symbolic elements which are transduced into the phonetic domain by some "phonetic implementation" function. Ambiguities about whether a certain descriptive generalization constitutes an aspect of the knowledge system (in the generative sense) – or, if it does, ambiguities about how that generalization is encoded and represented in the grammar – will by necessity make it much harder to settle disputes about whether marginal cases of typological gaps are genuine gaps or not. For example, a pair of segments in different languages which are phonetically (more or less) the "same" kind of segment may be quite distinct phonological objects in terms of their featural composition, with consequences for how such segments are predicted to pattern. A phonetically "voiced" stop is thus not necessarily a phonologically [voiced] segment (Jessen and Ringen 2002). Blevins (2006a) criticizes Kiparsky (2006) for explaining away some of her examples of final voicing on the grounds that they involve non-phonological voicing in this sense. To the extent that our model of the phonology-phonetics mapping is not wholly deterministic (and, in any case rather underdeveloped), such that we cannot at present unequivocally "discover" a language's featural representations from the phonetic signal, such criticism is not without justification. However, as long as the debate about intrinsic universals and their role in explaining the typology of sound patterns is taking place against the backdrop of a symbolic-generative model of grammar, ambiguities of this kind must be

taken seriously, and attempts at resolving them in a principled way must be sought.

The question whether diachronic-functional explanations are wholly sufficient in accounting for phonological typology, or whether they must be supplemented with (and/or replaced by) synchronic universals attributed to some version of UG, is far from settled. A somewhat independent question is the relative *intrinsic* merit of the diachronic and synchronic modes of explanation for linguistic phenomena. Blevins states, as the central premise of Evolutionary Phonology, that "[p]rincipled diachronic explanations for sound patterns *have priority over* competing synchronic explanations unless independent evidence demonstrates, beyond reasonable doubt, that a synchronic account is warranted" (Blevins 2004: 23, emphasis added; interestingly, in a later version (2006c: 124–125) downplays the synchrony/diachrony divide by substituting "phonological" vs. "extra-phonological" for "synchronic" vs. "diachronic"). This general idea is of course far from new (see Section 1), nor is it confined to phonology. In the domain of morphosyntax, for example, diachronic explanations of the typology of so-called "split ergative" systems (Anderson 1988; Garrett 1990) are widely accepted (e.g. Lightfoot 1999: 141; but see Kiparsky 2008b for counter-arguments). However, the territorial dispute between diachronic and synchronic/UG-based explanations has intensified sharply in the context of Optimality Theory in phonology, in which functionally "grounded" constraints are standardly proposed and attributed (usually) to an innate UG, as described in Section 2 above.

De Lacy (2006) suggests that there is no reason why the availability of a plausible diachronic-evolutionary explanation for some typological generalization should preclude the existence of a UG-internal explanation as well. While this is true at some level, the "priority" Blevins assigns to diachronic explanation is partly based on an Occam's Razor argument. Diachronic explanations of the sort advocated by Ohala, Blevins, and others are in a very fundamental sense *reductionist*: the explanandum (some recurrent sound pattern or typological generalization) is accounted for in terms of an explanans that is based in a concrete and observable domain which is subject to known physical laws and amenable to direct experimental verification (aspects of physiology, aerodynamics, acoustics, and perception). The alternative, to posit some innate constraint or constraint family as part of UG (or to rely on already-posited constraints in some novel ranking arrangement), locates the explanans in a domain which is itself hypothetical and essentially unobservable. With this in mind, it would seem to be sound methodology to operate under the assumption that nothing should be attributed to UG except when an adequate diachronic-functional explanation cannot be formulated (cf. Hale and Reiss 2000a). Of course, controversies such as that concerning the typology of voicing neutralization largely revolve around the question of when a diachronic account counts as fully "adequate" and when it does not.

Another claim made by de Lacy (2006) is that a theory like Evolutionary Phonology is about "performance," whereas generative theories are about "competence," and that since their domains are disjoint there is no inherent conflict and very little overlap. In practice, however, things are not so compartmentalized.

A central aim of diachronic-evolutionary models is to provide an account of cross-linguistic typology (including typological gaps and near-gaps), and typology-fitting has always been a major aspect of Optimality Theory as it is standardly practiced. By factorial typology, every proposal for a new constraint in UG is implicitly a claim about typological variation, and typological surveys are frequently the point of departure in optimality-theoretic analyses of specific phenomena. Be that as it may, it is certainly true that frameworks like Blevins' Evolutionary Phonology or Ohala's listener-based theory of sound change are not theories of the mental representation of synchronic-phonological knowledge (i.e. speakers' implicit knowledge of the sound patterns being explained). For this reason, one must take with a grain of salt Blevins' bold assertion that one consequence of Evolutionary Phonology is that "Markedness constraints are excised from synchronic grammars" (Blevins 2004: 23). As a technical term in Optimality Theory, a "Markedness constraint" merely refers to any constraint which evaluates the structural properties of phonological output representations. A parochial and functionally non-grounded constraint like the *ND constraint suggested by Hyman (2001a; see above) is no less an instance of a Markedness constraint than its functionally grounded near-opposite *NÇ. Likewise, the notion of Markedness constraints does not in itself presuppose innateness: even models which take constraints to be invented by the learner, rather than provided by an innate UG, rely on this construct (e.g. Hayes and Wilson 2008). What is at stake here is not the "excision" from synchronic grammars of ranked and violable output constraints as such, but rather a rejection of the general strategy of accounting for typological generalizations by positing universal, innate constraints on the sound shape of words.

8 Markedness and Universals as Learning Biases?

Many of the typological generalizations for which diachronic explanations have been successfully proposed are statistical tendencies. The pervasive recurrence of certain sound patterns reduces to recurrent sound changes with a clear phonetic basis, whereas the comparatively rare occurrence of certain other conceivable sound patterns (sometimes to the point of their being completely unattested) is attributed to the relative inaccessibility of these patterns via such recurrent types of sound changes. As rightly noted by some critics of diachronic approaches, standard generative models of phonology, such as Optimality Theory, have nothing to say about relative typological frequencies as such, since any constraint ranking permitted by UG is just as "good" as any other (de Lacy 2006). Nevertheless, some recent works have proposed an alternative conception of innate phonological knowledge, in the form of a set of *learning biases*, which do provide a means of dealing with typological asymmetries that are gradient rather than all-or-nothing.

In a series of recent studies, Moreton (2008a, 2010) attempts to quantify the relative robustness of the "phonetic precursors" to different types of sound patterns which have plausible diachronic origins in listener-based sound change, and

compares this robustness measure to the cross-linguistic frequency with which those same sound patterns are attested. Comparing patterns that relate tone to tone (e.g. of vowels in adjacent syllables) against ones that relate tone to laryngeal features (e.g. the voicing or voicelessness of an adjacent consonant), Moreton (2010) finds that even though the two have "precursors" – interactions and correlations in the acoustic signal – of approximately equal magnitude, phonological tone-tone interaction is vastly more common cross-linguistically than tone-voicing interaction. A similar case is made for height-height interaction vis-à-vis height-voicing interaction in Moreton (2008a): even though phonetic vowel height is affected more substantially by the voicing or voicelessness of an adjacent consonant than it is by the height of a vowel in a neighboring syllable, phonological vowel height harmony is much more common than phonological vowel raising/lowering conditioned by the voicing of adjacent consonants. On the assumption that the relative magnitude of a phonetic effect ought to determine (or at least be positively correlated with) the relative frequency of the resulting phonologized sound patterns, cross-dimension interactions (tone-voicing, height-voicing) thus seem to be "underphonologized" relative to same-dimension interactions (tone-tone, height-height).

This novel approach to the problem of phonologization is certainly very promising, but while Moreton's underphonologization argument appears compelling at first glance, it rests on certain assumptions which may be overly simplistic. Misperception (miscategorization, misparsing) is the key process in listener-based models of sound change, in that the listener either overcorrects or undercorrects for a contextual effect (be it real or apparent) in the acoustic signal. However, it is far from obvious that a direct parallel can be drawn between some raw acoustic-auditory measure of the "magnitude" of an interaction on the one hand and how that interaction is going to be dealt with by the listener in the perception/parsing process on the other. The kinds of phenomena that Moreton deals with would all fall under Ohala's hypocorrection rubric, whereby coarticulatory compensation (Beddor and Krakow 1999; Beddor *et al.* 2002) is underapplied, such that a coarticulatory effect is misinterpreted as an intrinsic property of the affected segment. Ohala (1994a) has suggested that listeners may be more attuned to – and hence more likely to compensate successfully for – the influence of stronger coarticulation triggers than that of weaker triggers. For example, this would explain why in most front/back vowel harmony systems, it is precisely those vowels which ought to be the strongest inducers of coarticulatory fronting on neighboring vowels (namely /i, e/) that are instead "neutral," failing to trigger any phonological fronting at all. If Ohala's (1994a) theory of asymmetric coarticulatory compensation has general validity, this would predict that interactions with greater "magnitude" (more robust phonetic precursors, in Moreton's sense) ought in fact to be *less* likely to become phonologized. Needless to say, this is exactly what the typological data show; the case for "underphonologization" may thus turn out to be illusory.

Such caveats aside, a second and no less important aspect of Moreton's work is his demonstration that artificial language-learning experiments yield results

which do mirror the typological distribution closely: cross-dimension patterns (height-voicing interaction) are learned far less reliably than same-dimension patterns (Moreton 2008a).[8] In this particular case, the factor which governs ease of learning can be construed as the inherent complexity of the interaction pattern (cf. Pycha, Nowak, Shin, and Shosted 2003; Wilson 2003). Moreton (2010) accounts for the supposed underphonologization of cross-dimension interactions in terms which go beyond superficial measures of formal simplicity. He proposes an explicit Bayesian learning algorithm (set within an Optimality Theory framework) in which the language learner is biased against adding a new constraint – that is, inventing or constructing it on the basis of input data – if that constraint interacts with a large number of other constraints that are already present in the grammar (see also Moreton 2008b).

In a similar vein, Wilson (2006a) reports asymmetries in artificial learning tasks where the determining factor cannot be construed as formal simplicity but rather revolves around "phonetic naturalness" in some sense. Subjects who learned a velar palatalization pattern [ke] → [tʃe] tended to generalize the [k] → [tʃ] change to [ki] contexts as well, whereas subjects who learned [ki] → [tʃi] did not generalize this to [ke] contexts. Wilson (2006a) argues for a conception of markedness and phonetic naturalness as a "substantive bias" on phonological learning. Differential treatment of equally-unattested onset clusters is another area which has recently been claimed to show evidence for such substantive bias, namely the greater "splittability" of obstruent-obstruent clusters like /bd/ than of obstruent-sonorant clusters like /bn/ in infixation (Zuraw 2007 on Tagalog) and in epenthesis (Berent, Steriade, Lennertz, and Vaknin 2007 on English). However, such findings are often hard to interpret unequivocally. For example, even though neither /bn/ nor /bd/ occur as onsets in English, they fit into larger-scale categories which are attested (stop + sonorant) and unattested (non-/s/ obstruent + obstruent), respectively (Albright 2009).

Another place where substantive constraints may be in evidence is in the diachronic development of the lexicon. Frisch, Pierrehumbert, and Broe (2004) suggest that the elaborate (gradient) dissimilatory cooccurrence restrictions on consonantal roots seen in Semitic languages have evolved as a result of the cognitive pressures affecting the lexicon incrementally: "lexical items that avoid repetition will be easier to process, and so will be favored in acquisition, lexical borrowing, coining novel forms, and in active usage" (Frisch *et al.* 2004: 221). In an intriguing study charting lexical development from Proto-Indo-European via Latin to Modern French, Martin (2006, 2007) demonstrates how the cross-linguistically typical frequency relation between /b/ (more frequent) and /d/ (less frequent) has gradually manifested itself over time. Not only do /b/-initial lexical items show a greater "survival rate" than /d/-initial ones, they also appear to be favored in the formation of new words through morphological derivation, as well as in borrowing. For example, even though /d/-initial words outnumbered /b/-initial ones both in the Latin lexicon and in that of Classical Greek – largely due to the near-absence of /b/-initial words in Proto-Indo-European – Latin borrowed far more /b/-initial than /d/-initial words from Greek.

The alternative view that universal "markedness" or "naturalness" takes the form of substantive biases on the acquisition of sound patterns, and perhaps on language use, is by no means incompatible with diachronic-functional approaches. For example, if such biases do exist, they would simply be one of the many potential sources of "external error" in evolutionary simulation models like that of Wedel (2007). Blevins (2006a) stresses that her Evolutionary Phonology by no means rules out the potential existence of innate knowledge or cognitive processing effects. However, to the extent that substantive biases exist which duplicate recurrent sound changes with plausible sources in misperception (e.g. the greater propensity for [k] > [tʃ] to take place before [i] than before [e]), this does most certainly complicate the task of teasing apart the true explanatory factors which underlie specific typological generalizations. More research in this area is clearly needed in order to shed further light on this contentious issue.

9 Summary

What role to attribute to diachronic change in explaining the typology of sound patterns, the characteristic phonetic naturalness of such patterns, and the various parochial details of individual sound systems, is a question which is as old as the discipline of phonology itself. It is probably safe to say, however, that the current debate surrounding the role of diachronic explanation in phonology, and the relative importance of this and other types of explanatory factors, is more active than it has ever been in the history of modern phonological theory. This has partly been brought on by the formulation of elaborate and explicit theories of sound change. In addition, our knowledge of the full range of typological variation in the sound systems of the world's languages has become much more principled and more detailed – helped along in large part by the inherently typological orientation of the Optimality Theory enterprise – and this has revealed just how strongly the patterns of cross-linguistic variation seem to reflect phonetic constraints on speech production and perception. Finally, a tremendous growth in experimental and computational methodologies and approaches to phonological problems has shifted the arena of this debate out of the armchair and into laboratories (real and virtual alike). It will be interesting to see to what degree the mainstream of generative phonological theory will be shaped by these developments over the next decade or two.

ACKNOWLEDGMENTS

The work underlying this article was partly supported by research grant 410-2004-0710 from the Social Sciences and Humanities Research Council of Canada. This chapter is reprinted from *Language and Linguistics Compass* 2(5), 859–893 (2008), with kind permission from Wiley-Blackwell and the journal editors. Some bibliographic citations have been

updated, but no substantive changes or additions to the text have been made. The work underlying this article was partly supported by research grant 410-2004-0710 from the Social Sciences and Humanities Research Council of Canada.

NOTES

1 As the square brackets around [B] suggest, we are here dealing with "B" as an element that occurs in phonological *surface* representations in the language. Whether such a pre-existing [B] reflects an independent phoneme /B/ in *underlying* representations is essentially an orthogonal issue; in principle, [B] might just as well be a predictable surface variant of some other phoneme (it might even be conceivable for [A] and [B] to constitute allophones of the same phoneme). In any case, Ohala's model of sound change does not necessarily presuppose the kind of underlying vs. surface representational distinction assumed in generative models. It should also be kept in mind that "A" and "B" need not represent individual segments. In the English example cited above (/swoɹd/ > /soɹd/, etc.), a bisegmental C+[w] cluster (= "A") is being misperceived, hypercorrectively, as monosegmental C (= "B") plus coarticulatory rounding from a neighboring vowel.

2 Such asymmetric confusion is unremarkable in and of itself, and is well attested in other perceptual domains. For example, in visual letter recognition, one is far more likely to misperceive a "Q" as "O" (by failing to notice the distinctive tail that sets the former apart from the latter) than to misperceive "O" as "Q" (by spuriously introducing that same characteristic feature). Plauché *et al.* (1997) suggest that the /ki/ vs. /tʃi/ asymmetry is due to precisely such an all-or-nothing feature in the acoustic signal, namely the spectral peak in the 3–4 kHz range that is characteristic of velars. Guion (1998) is skeptical of this explanation, conjecturing instead that the asymmetry is likely due to palato-alveolar affricates being a perceptually more robust category than velars, owing to their longer duration and greater amplitude.

3 In an explicit reference to exemplar-based models (see Section 6), Blevins uses the terms "best exemplar" or "prototype" in this context, even though such models typically do not invoke such abstractions at all. On this point, Blevins's proposal most closely echoes Bybee (2001), who advocates a kind of hybrid between an exemplar model and a prototype model – though even Bybee does not appear to presuppose that each word form is somehow characterized by a *single* ("prototypical") phonological representation in the lexicon.

4 Though he does not explicitly acknowledge this parallel, Silverman's (2006a, 2006b) account of Trique labialization is essentially identical to Pierrehumbert's (2002) suggested explanation for the hyperarticulation of words with many lexical neighbors (see Section 4 for discussion). In both cases, more "exaggerated" tokens are more likely to be categorized correctly, and will thus gradually – through accumulation over successive cycles of speaker-listener/learner interactions – end up being overrepresented in the exemplar cloud for the category in question.

5 However, it seems that this account does not provide an explanation for why the spread of [h] variants should target (phrase-medial) __#V contexts specifically, to a much greater extent than it did (phrase-final) prepausal contexts.

6 Zsiga, Gouskova, and Tlale (2006) cite conflicting phonetic evidence and call into question Hyman's (2001a) analysis of Tswana in terms of a *ND constraint. However, the

existence of a postnasal devoicing process in Tswana is confirmed by Coetzee, Lin, and Pretorius (2007).

7 In the hypothetical Gujarati-based example that Kiparsky (2008b) constructs, this is in fact rather trivially obvious. Given that [ə] already exists in the language, the [a] > [ə] merger would result in the language having two kinds of [ə], one stress-repelling and one stress-attracting. An antepenultimate or final [ə] would thus attract stress away from a mid or high penultimate vowel in some words but not others. Similarly, a mid or high vowel in the antepenultimate would attract stress away from a penultimate [ə] in some words but not others. This unfortunate flaw in the particular example chosen does not detract from what is almost certainly the main point of Kiparsky's argument. We might as well imagine a sonority-driven stress system where the relevant scale is [a] > [e, o] > [i, u], with no [ə] vowel in the language; a sound change [a] > [ə] would indeed turn this into a system in which [ə] attracts stress over and beyond more sonorous vowels like [o] or [e]. Presumably Kiparsky would expect this, too, to result in lexicalization of stress in the hands of subsequent generations of learners, in spite of the fact that stress placement would still be completely predictable, observationally speaking.

8 Interestingly, the more easily learned same-dimension patterns include not only vowel height harmony but also (long-distance) consonant voicing harmony, which is typologically quite rare in comparison (Hansson 2010) and may be largely dependent on complex diachronic scenarios for its emergence (Hansson 2004a).

11 Phonetics in Phonology

D. R. LADD

1 Introduction

The primary semiotic medium of spoken language consists of acoustic signals – sound waves – produced by articulatory gestures in the human vocal tract and processed by human auditory systems. To understand more about this aspect of language it would therefore seem scientifically appropriate, even necessary, to learn more about the human vocal tract, the human auditory system, and the physics of sound. At the same time, it has been clear for more than a century that language uses the medium of sound in a very specific way, which involves the human cognitive capacity for creating categories and symbolic systems. This capacity makes it possible for two physical (acoustic) events that are objectively quite different to count as instances of the same category in the symbolic system, and for two physical events that are objectively very similar to count as instances of two different categories. It also makes it possible for different languages to categorize the physical events of speech in different ways. If we want to understand the medium of spoken language, therefore, it is not enough to consider only the physical aspects of the production, transmission, and perception of sound; we need to consider the symbolic value of the sounds of speech as well.

The dual nature of speech sounds – as physical events and as elements of a symbolic system – has been recognized since the emergence of the phonemic principle in the late nineteenth century; in some sense, the emergence of the phonemic principle and the recognition of the dual nature of speech sounds were one and the same scientific achievement. Since the 1930s, and especially since Trubetzkoy's *Principles* (1958 [1939]), it has been customary to reserve the term *phonetics* for the study of the physical aspects of speech sounds – what Trubetzkoy described as "the study of the sounds of [Saussurean] *parole*" – and to use the

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

newer term *phonology* for "the study of the sounds of *langue*" (Trubetzkoy 1939: 7).[1] This terminological distinction is now such a fundamental part of our conceptual landscape that it seems perfectly normal for the editors of a volume on phonological theory to solicit a chapter on "phonetics in phonology." At the same time, the need for such a chapter shows that the distinction itself continues to engender conceptual difficulty. It is fairly obvious what "the sounds of *parole*" might refer to, but less obvious what "the sounds of *langue*" might be. Understanding the relation between phonetics and phonology is thus ultimately a matter of understanding the dual nature of the sign, and much of the difficulty in defining and delimiting their respective realms is ultimately due to the difficulty of deciding what sort of abstractions we are dealing with when we study language.

In the long run, the broader task of what we might call the phonetic sciences is to understand the human capacity for categorizing the sounds of speech, and to understand how this capacity reflects – and is reflected in – the structure of language. In this chapter I take some such unified ultimate goal for granted. I realize that not everyone would subscribe to it in the form in which I just stated it, and in any case there are plenty of challenging subsidiary questions on both the physical side and the symbolic side to keep researchers fully occupied without thinking about long-term goals. However, I hope to demonstrate that phonetics and phonology are inextricably intertwined even in theories that purport to draw a sharp distinction between them, and that the place of phonetics in phonology has been absolutely central ever since the phonemic principle emerged. In particular, I aim to show that many standard concepts in phonology depend crucially on the body of theory and practice that we can refer to as *systematic phonetics*. That is, most twentieth-century phonology – the study of the sounds of *langue* – is based firmly on a theory of phonetics – the sounds of *parole.* To the extent that there are problems with the theory of phonetics, therefore, there are problems with phonology, and this chapter also attempts to outline what some of those problems are and how we might integrate an empirically more defensible view of phonetics into our understanding of phonology.

2 Systematic Phonetics in Phonology

The term "systematic phonetics" is apparently due to Chomsky (1964: Chapter 4), but the idea of systematic phonetics is embodied in the principles of the International Phonetic Association (IPA[2]). These principles are stated in summary form in successive editions of the *IPA Handbook*, and are discussed at greater length in the handbook's most recent edition (IPA 1999) and in textbook presentations of phonetics such as Laver (1994). Systematic phonetics depends on two key premises, which I will refer to as the *segmental idealization* and the *universal categorization assumption*. These may be stated as follows:

The segmental idealization: speech (NB not language) can appropriately be idealized as a string of ordered discrete sound segments of unspecified duration. ("Phonetic

analysis is based on the crucial premise that it is possible to describe speech in terms of a sequence of segments" (IPA 1999: 5).)

The universal categorization assumption: there is a closed universal inventory of possible segment types ("The IPA is intended to be a set of symbols for representing all the possible sounds of the world's languages" (IPA 1999: 159)).

These premises were incorporated largely without comment into virtually all theorizing about phonology from the 1940s until the 1990s and are still widely accepted. Together, they yield the key theoretical construct generally known as the *phone*, and, as a kind of corollary, the notion of *(distinctive) feature.* These ideas are now such a fundamental part of the way we think about phonetics that it comes as a surprise to realize that they were not taken for granted until well into the twentieth century, and it is worth taking some time to trace their development.

2.1 The Phone

The phone has been part of the IPA enterprise from the very beginning, but at first it was only implicit. According to the history of the International Phonetic Association included in the *IPA Handbook* (IPA 1999: 194–197), the IPA started out life in 1886 as, in effect, a response to the inconsistencies of English orthography, aiming at a practical orthography with consistent phoneme-grapheme correspondences for use in language teaching. However, the idea of developing a consistent practical orthography adaptable to all languages was explored very early in the history of the IPA, and the first version of the IPA alphabet was published in 1888, along with a set of principles on which it was based. The first of these principles (again, according to IPA 1999) was:

> There should be a separate sign for each distinctive sound; that is, for each sound which, being used instead of another, in the same language, can change the meaning of a word.

In modern terms, this very clearly states that IPA transcription is intended as a phonemic transcription, and *sound* is clearly being used to mean "phoneme." However, the seeds of theoretical confusion were sown immediately, in the second principle:

> When any sound is found in several languages, the same sign should be used in all. This applies also to very similar shades of sound.

This second principle requires us to define "sound" in a different way from the first principle, because we cannot use any sort of practical test based on word meaning to decide whether two sounds in two different languages are distinctive or not. The notion of sound in the first principle is language-specific; the notion of sound in the second implies a language-independent categorization. This second sense of sound is what came to be known as the phone.

Leonard Bloomfield, the central figure in American linguistics in the first half of the twentieth century, saw the contradiction between these two principles and devoted several pages of his great work *Language* (1933) to trying to expose it. He uses the term "phonetic [sic] transcription" to refer to "a system of written symbols which provides one sign for each phoneme of the language we are recording," and explicitly denies the validity of attempts to transcribe non-distinctive acoustic detail. It is worth quoting him at some length:

> Having learned to discriminate many kinds of sounds, the phonetician may turn to some language, new or familiar, and insist upon recording all the distinctions he has learned to discriminate, even when in this language they are non-distinctive and have no bearing whatever. . . . The chief objection to this procedure is its inconsistency. The phonetician's equipment is personal and accidental; he hears those acoustic features which are discriminated in the languages he has observed. Even his most 'exact' record is bound to ignore innumerable non-distinctive features of sound; the ones that appear in it are selected by accidental and personal factors. . . . [H]is most elaborate account cannot remotely approach the value of a mechanical record.
>
> Only two kinds of linguistic records are scientifically relevant. One is a mechanical record of the gross acoustic features, such as is produced in the phonetics laboratory. The other is a record in terms of phonemes, ignoring all features that are not distinctive in the language. . . . (pp. 84–85)

However, Bloomfield's views had essentially no influence on subsequent theoretical developments, not even among his closest followers, the so-called neo-Bloomfieldians like Bernard Bloch (e.g. 1941, 1948) and Charles Hockett (e.g. 1942, 1955). Instead, the idea that there is a valid universal basis for abstracting segment-sized sounds out of the stream of speech, and a valid universal framework for categorizing them, became firmly established in the 1920s and 1930s.

It is true that there was at least one attempt to put the phone idealization on a firm theoretical footing. In his (1943) monograph *Phonetics*, Kenneth Pike devoted an entire chapter (entitled "Units of sound") to the theoretical difficulties with the notion "speech sound" or "phone," stating the problem as follows:

> Speech, as phoneticians well agree, consists of continuous streams of sound within breath groups; neither sounds nor words are separated consistently from one another by pauses, but have to be abstracted from the continuum. Phonemicists concur in the belief that some unit of speech, the phoneme, can be discovered as the basic constituent of a linguistic system. . . . Is there a significant halfway point between the continuum and the phoneme? Is there a real, nonfictitious segment of sound which is not a phonemic one? (p. 42)

Bloomfield's answer to Pike's question, as we just saw, was unambiguously "No." Pike, however, after some discussion of differing views and difficult cases, inclines toward "the conclusion that there must ultimately be some such phonetic segmentation behind speech" (p. 46). He then sets out (p. 52) to find "a workable method for the delineation of natural phonetic segmentation," in which the "segmental unit is to be determined entirely apart from phonemic function." He notes that "[a] corollary of this aim states that such a segmentation procedure

is equally applicable to any and all languages, or to any stream of nonsense syllables." Such a procedure means that "an impressionistic phonetic record of a new language proves *theoretically legitimate* as well as practically valuable . . . for the phonemicist . . ." (p. 53, emphasis added).

It is difficult to know to what extent Pike's theoretical considerations influenced the development of the field, but it is clear that few writers after him were worried about the theoretical legitimacy of the phone idealization, or about the assumption that there is a closed universal set of phones. By 1949 (when the IPA *Principles* were republished in revised form), the notion of discrete speech sounds or phones appears to be taken for granted. The new version of the first principle starts: "When two sounds occurring in a given language are employed for distinguishing one word from another . . ."; and the second begins: "When two sounds are so near together acoustically that there is no likelihood of their being employed in any language for distinguishing words . . ." (IPA 1949: 1).

Pike's reference to the practical value of his procedures for the "phonemicist" reminds us of the central role that systematic phonetics had already come to play in theoretical phonology. During the 1930s and 1940s the phone idealization became firmly embedded in linguistic discussions of the phoneme on both sides of the Atlantic – as for example in Trubetzkoy's discussion of how to define and identify phonemes (1939: Chapter II), which simply presupposes the phone (*Lautgebilde*, translated as *sound* by Baltaxe; cf. her translator's note (1969: 36)). Early codifications of the "allophone" idea (e.g. Bloch 1941; Hockett 1942; cf. Trubetzkoy's "combinatory variant") are probably the clearest illustration of the central importance of the phone concept in shaping phonological theory.

Consider the realization of voiceless stops in English syllable onsets, which is probably used as an example in 90% of beginning linguistics courses in the English-speaking world. It is well known that in absolute initial position, as in *peach*, voiceless stops typically have a voice onset time (VOT) in the general range of 50–70 ms, whereas when preceded by /s/ in an onset cluster, as in *speech*, they typically have a VOT in the general range of 0–20 ms. This is an easily observable fact about the phonology of English, and provides a clear and simple illustration of the fundamental phonological concept of lawful conditioned variation. However, statements of this variation are conventionally expressed not in terms of mean VOT, but in terms of two phones, usually notated (for example) [p] and [p^h], the latter occurring in absolute initial position and the former occurring after /s/. This statement is already a considerable abstraction away from observations about VOT, but that is not acknowledged in most classical formulations of the phoneme or in most textbook presentations. Instead, the phones are considered to be the *raw data*; transcriptions like [spitʃ] and [p^hitʃ] are assumed to provide a faithful representation of what a speaker really produces. Rather than recognize [p] and [p^h] as abstractions based (as Bloomfield emphasized) on the personal equipment of the transcriber,[3] classical phoneme theory took them as categories of phonetic description, identifiable in a language-independent way.

I am, of course, well aware of operating with the benefit of hindsight here. When I say that the facts about English VOT are "easily observable," I am referring

to the technological environment of today, not that of the 1950s or even the 1980s. Today, given free software like Praat (http://www.praat.org) and readily accessible tools for plotting data, it is indeed a simple matter to establish that the facts of English VOT are roughly as I have stated them and to see clearly that such facts are a matter of statistical distributions, not unvarying categories. However, an early attempt to base phonology on such a statistical view of phonetics (Zwirner's "Phonometrie"; Zwirner and Zwirner 1966 [Zwirner 1936]) was rejected by Trubetzkoy (1939: 10–12) in what may be seen as an early instance of the gulf of misunderstanding between phoneticians and phonologists. Even after the sound spectrograph brought general awareness of the variability of the raw data – triggering considerable soul-searching at least on the part of neo-Bloomfieldian phonologists (see e.g. Joos 1948; Bloch 1948, especially footnote 6 and Postulates 9 and 11; Hockett 1955, Section 5) – the phone idealization always managed to survive.[4]

The supposed reality of phones was crucial to the role played in traditional definitions of the phoneme by the minimal pair test, that is, the substitution of one sound for another. Postulating a phonemic distinction between /p/ and /b/ in English depends in part on agreeing in advance that [p^h], [p], and [b] are comparable sounds or segments in pairs like *pit/bit, pang/bang, cap/cab, poor/boor*, and so on. In the case of [p^h], [p], and [b], there is little disagreement that these are comparable units, but there are many well-known cases where there was no such agreement and phonemic analysis was correspondingly controversial. The best-known case in English is probably that of the affricates, and the problem of whether to treat affricates and other such complex segments as single phonemes or as clusters has a long history. The relevance of segmentation to these cases is as follows: if *chip* begins with the phones [t] and [ʃ], then [t] can be replaced by zero to yield *ship* and [ʃ] by [ɹ] to yield *trip*, so that *chip* can be said to begin with a cluster; if, on the other hand, we do not identify the first part of the affricate with the phone [t] and/or do not identify the second part with [ʃ], then there is no obstacle to treating the affricate as one phone and analyzing the beginning of *chip* as a single consonant. Without a universally valid method of segmentation *and a universally valid system of classifying segments as the same or different*, defining phonemes in terms of the distribution of phones is ultimately arbitrary, as Pike correctly saw. Pike's faith that such a segmentation could be justified theoretically was not shared by for example, Martinet (1966 [1939]), who says: "From all this, it turns out that the first task of the phonologist is an in-depth phonetic analysis of the language under study, during which analysis it will be necessary above all to be careful not to be led astray by the imperfections of traditional phonetic transcriptions" (p. 122, my translation). In other words, Martinet recognizes that the identification of the phones on which we base our theoretical definition of the phoneme is specific to a given language.

Nevertheless, twentieth-century theories of phonology were universally built on the assumption that phones and phonetic transcriptions are a scientifically appropriate language-independent representation of speech. This was the idea that Chomsky picked up in his brilliant dissection (1964) of mid-century phoneme theory

and his presentation of the assumptions underlying what became mainstream generative phonology. He drew a sharp distinction between "physical phonetics" and "systematic phonetics," explicitly claiming that both levels of description are necessary in a formal model of language and speech. Specifically, he envisaged an overall theoretical structure in which the output of the phonology (or, more broadly, the output of the grammar) is a systematic phonetic representation consisting primarily of a string of phones; this systematic phonetic representation is then passed to a phonetic implementation system – not part of *langue* – where universal biomechanical and physical principles generate the physical phonetic output. In terms that have become familiar more recently, generative phonology thus sees the systematic phonetic representation as the *interface* between phonology and phonetics – or, if we accept Trubetzkoy's definitions, the boundary between *langue* and *parole*. As is well known, Chomsky argued that the "taxonomic phonemic" level of the neo-Bloomfieldians was unnecessary and unmotivated, and that the phonological grammar should map directly from abstract "systematic phonemic" representations to the systematic phonetic output (cf. also Halle 1959). Like the neo-Bloomfieldians, however, he did not question the assumption that the systematic phonetic representation is a scientifically valid idealization. Indeed, this assumption was vigorously defended by Postal (1968) and with very little further discussion was incorporated into the generative theory codified in *SPE* (Chomsky and Halle 1968) and a number of textbook presentations in the 1970s (e.g. Schane 1971a; Hyman 1975; Kenstowicz and Kisseberth 1979).

Since the 1960s, few phonologists have questioned the early generative acceptance of systematic phonetics and the segmental idealization, and the idea of universal phonetic categorization remains at the foundation of most present-day work in phonology. It is true that in the late 1980s and early 1990s there was a flurry of interest in interface issues. In 1990 the question of phonetic representation occupied an entire special issue of the *Journal of Phonetics* (vol. 18: 297–477), in which the segmental idealization was attacked (e.g. Pierrehumbert 1990), assumed (e.g. Ladefoged 1990), and defended with empirical evidence (e.g. Nearey 1990). However, at more or less the same time the attention of the field was captured by Optimality Theory (OT; e.g. Prince and Smolensky 2004; Archangeli and Langendoen 1997; Kager 1999) and interface issues were largely marginalized. OT incorporates the generative understanding of phonetics wholesale: its entire architecture is based on having a set of categorically distinct "outputs" to evaluate, which is possible only if we abstract away from the infinite variability of speech and assume some sort of universal categorization of the speech sounds. Moreover, the key faithfulness constraints with which the theory began, Parse and Fill (and their successors Max and Dep), are built on the assumption that the output can be exhaustively and unambiguously divided into segments. Within OT, there have been some attempts to deal with the empirical difficulties posed by these assumptions (notably Boersma 1998), but the great body of work in OT continues to accept systematic phonetics as a valid basis for describing the output of the grammar, and as a convenient delineation of the boundary between its concerns and those of others.

2.2 Distinctive Features

The idea of a universal scheme of classification for phones gives rise to what is perhaps the central theoretical construct of mid-twentieth-century phonology, namely the feature. In an informal way, of course, the dimensions of the IPA symbol chart are a kind of feature analysis, but we are concerned here with the place of such classification *in phonology*. Linguists had long been aware that certain kinds of sound changes are common and somehow natural, and that common phoneme inventories across languages are often quite symmetrical if described in terms of phonetic dimensions. But this awareness played no formal role in most Anglo-American phonemic theorizing, which was almost exclusively concerned with the procedures for grouping phones into phonemes. The work of putting phonetic symmetries and similarities on an explicitly phonological footing was carried out by the members of the Prague School during the 1930s.

The basic ideas were presented by Trubetzkoy in *Principles*. Trubetzkoy's theoretical starting points were, first, the strict separation of phonetics and phonology, and second, the structuralist or Saussurean idea that language involves a system of oppositions, in which the central property of any given sign is that it is *not* any of the other signs. This last idea is the view summed up in Saussure's well-known dictum "Dans la langue il n'y a que des différences" and in Jakobson and Halle's suggestion (1956: 22) that the meaning of a phoneme is "mere otherness." Accordingly, Trubetzkoy starts out by describing phonology in purely abstract terms: "The signifier of the system of language [i.e. of *langue*] consists of a number of elements [viz., phonemes – DRL], *whose essential function it is to distinguish themselves from each other*" (Baltaxe 1969: 10, emphasis added)). Nevertheless, in order to talk about the actual systematic differences that distinguish one phoneme from another – differences in *langue* – Trubetzkoy did not refer to abstract dimensions but to concrete phonetic properties of phones – elements of *parole*. He treats this recourse to phonetic dimensions as inevitable: "As regards phonology, it is clear that it must make use of certain phonetic concepts. For instance, the claim that in Russian the contrast between voiced and voiceless obstruents is used to differentiate between words belongs to the field of phonology. The terms 'voiced' and 'voiceless' and 'obstruents' themselves, however, are actually phonetic" (1969: 14). He reiterates the necessity of making this link to phonetic concepts at greater length in the introduction to Chapter IV (1969: 91ff.).

Trubetzkoy's version of features (or "oppositions") was thus in some important respects merely an expedient codification of the dimensions of the IPA chart. In particular, the distinction he draws among "privative," "gradual," and "equipollent" oppositions is patently related to – if not actually influenced by – the physical nature of those dimensions, and much of his discussion is cast in traditional IPA terms. However, three major subsequent developments meant that the feature concept took on a theoretical life of its own.

The first development was the publication of Jakobson, Fant, and Halle's *Preliminaries to Speech Analysis* (1952; henceforth *JFH*), which presented a fully

worked-out theory of distinctive features whose dimensions were no longer merely those of the IPA. The most conspicuous taxonomic innovations were that the features were exclusively binary and that they were defined in purely acoustic terms. However, the *JFH* feature system reaffirms the two premises of systematic phonetics identified at the beginning of this discussion: it presupposes the segment, and it explicitly presents the taxonomic framework as universally valid. Actually, the *JFH* version of the segmental idealization does represent a refinement of the IPA version, because it acknowledges the continuous variation of acoustic parameters. Specifically, it treats the segment not as a section of signal with duration, but as an idealized instantaneous slice through the signal at a specific point in time: "For practical purposes each phoneme can be represented by a quasi-stationary spectrum in which the transfer function is invariable with respect to time . . ." (1952: 18). The features that characterize the segment are therefore based on the acoustic properties of the signal at the point in time when the idealized instantaneous slice is taken. Apart from that refinement, however, the *JFH* approach is built on a conception of the phone that was perfectly consistent with the ideas of Pike or Hockett.

The *JFH* definition of segment, and the concomitant definition of the feature as an actual acoustic property at an identifiable point in time, is part of a second important development in feature theory that is much less widely recognized. For Trubetzkoy, features are above all *abstract* characteristics of phonemes: phonemes are the elements of phonology, forming part of a system of *oppositions*, and phonetic properties are of interest only insofar as they describe how the abstract oppositions are manifested. The proposal in *JFH* that phonemes are instantaneous time slices at which features can be identified *in the signal* represents a considerable departure, in that the features have become acoustic events or properties of acoustic events rather than abstract dimensions. This in turn easily leads to the idea that the elements of phonology are features, and phonemes are composite. Such a conception is strongly suggested by *JFH* and made explicit by Chomsky and Halle's work in the 1950s and 1960s, but is clearly absent from Trubetzkoy's thinking.[5]

This finally brings us to the third major development of the feature notion, namely its incorporation into the phonological theory of *SPE*. In some respects the *SPE* version of feature theory was conservative: it did not question the assumption that features should provide a universal framework for describing actual sounds, and it did not pursue the *JFH* definition of the segment as an instantaneous time-slice, conceiving of sounds very traditionally as phones. However, it formally adopted the notion that features are the primitive elements of phonology and phonemes merely sets or "bundles" of such primitives. Moreover, it took seriously another idea, implicit in Trubetzkoy but not developed in *JFH*, namely that the universal descriptive framework established by the set of features should also allow us to express phonological symmetries and generalizations. This led to the better-known aspect of Chomsky and Halle's revision of *JFH*, namely the replacement of several of the acoustically-based *JFH* features such as [grave] and [compact] by features based on articulatory dimensions more like the traditional

dimensions of the IPA chart. The principal justification for these changes was that the new features were better suited to expressing the generalizations of phonology. Like Trubetzkoy, that is, Chomsky and Halle seem to have concluded that the best way to give a description of phonological regularities was in terms of the taxonomic dimensions of phonetics.

Considering the importance that Trubetzkoy attached to the phonology-phonetics distinction, the persistence of traditional phonetic dimensions in phonology is striking. One could perfectly well imagine a description of the distinctive oppositions in a given language that makes no reference to phonetics and really does work with the idea of abstract distinctness or "mere otherness." Standard names for the four tonemes of Mandarin Chinese are essentially of this sort: the long-standing Western practice of using the numbers 1 to 4 obviously makes no reference to the pitch contours by which the abstract tonemes are phonetically manifested. (Essentially the same is now true of the traditional Chinese names *yīn píng* 'yin level', *yáng píng* 'yang level', *shàng* 'upper', *qù* 'leaving', though in Classical Chinese these may have had some phonetic content.) Indeed, this might seem to be a good way of pursuing Trubetzkoy's professed goal of categorizing "the sounds of *langue*": such names or numbers are shorthand ways of referring to abstract phonological elements that are functionally equivalent across the lexicon irrespective of phonetic realization. For example, "Tone 2" is mid-high-rising in standard Mandarin and mid-low-falling in Chengdu (Chang 1958). The phonetic realization could hardly be more different, but the *system* of tones in both varieties is still basically the Mandarin four-tone system, in the sense that words having "Tone 2" in one variety will reliably have it in the other as well.

It is true that non-phonetic names like "Tone 2" are names for whole phonemes, not features, but there is no obvious reason why non-phonetic names could not also be used to designate the patterns of opposition that Trubetzkoy saw as the essence of phonology. Indeed, it is not hard to see that phonetically abstract names for phonologically relevant dimensions are sometimes exactly what we want. Perhaps the clearest instance is Chomsky and Halle's proposal for a feature [syllabic] to replace the *JFH* feature [vocalic]. Although they provide an ostensibly phonetic definition of [syllabic] as "constituting a syllable peak" (1968: 354), they give little indication of the difficulty of defining syllables phonetically, and the motivations for having such a feature are patently phonological. Similar remarks could be made about the feature [tense] applied to vowels in English or Dutch, or about the descriptive term *rhotic*, which is sometime used to refer to the phonetically diverse set of segment types that manifest the /r/ phoneme in English and other European languages.

Nevertheless, the unquestionable descriptive utility of such phonetically abstract features has not so far raised any serious theoretical doubts about the appropriateness of using phonetic dimensions to characterize phonological oppositions. On the contrary, a good deal of theoretical work (e.g. Hayes and Steriade 2004) has examined the "grounding" of phonological features in phonetics, and the phonetic basis of feature definitions is now seen as involving a significant

theoretical claim, "namely, that natural phonological classes and sound changes will be definable in phonetic terms" (Kenstowicz and Kisseberth 1979: 240). Following Postal 1968, Kenstowicz and Kisseberth refer to this claim as the "naturalness condition" and assume its validity. For example, they say explicitly of the feature [syllabic] that "[s]ince the syllable has not yet been defined satisfactorily in phonetic terms, the phonetic correlates of this feature are unclear" (1979: 242), implicitly presupposing that such satisfactory phonetic definition will eventually be forthcoming. This presupposition is made explicit when they note more generally that "there are still a number of widespread phonological processes which presuppose natural classes of sounds for which no straightforward phonetic correlates are presently known. They pose a challenge to future research and one can only hope that as phonetic science progresses, these unexplained counter-examples to the naturalness condition will eventually be resolved" (1979: 241). In short, they treat any difficulties in reconciling phonetic and phonological uses of features as a matter for empirical research rather than theoretical reconsideration.

3 Systematic Phonetics in its Own Right

In the discussion so far I have sought to show that a crucial component of most contemporary conceptions of phonology is a theory *of phonetics*: the rigid separation between phonetics and phonology posited by Trubetzkoy and assumed by subsequent generations of linguists is illusory (cf. also Chomsky 1964: 109f [1972: 423]). The illusion could be harmless, of course. As long as the theory of phonetics is approximately valid, then what I have said so far amounts to little more than an academic exercise in the exegesis of classic texts. That is, it could be that Trubetzkoy was wrong about the strict division, but nothing else of substance changes. In this case, Kenstowicz and Kisseberth would be justified in awaiting the results of further empirical progress in phonetic science.

However, since the 1980s progress in phonetic science has been considerable. The increasing ease of acquiring instrumental data – especially acoustic data, but also articulatory data – means that we know more and more about the details of phonetic realization. Much of this research has been carried out under the heading of "laboratory phonology" (e.g. Kingston and Beckman 1990; Pierrehumbert, Beckman, and Ladd 2000), a phrase that would probably have struck Trubetzkoy as an oxymoron. But the phrase is precise and meaningful: laboratory phonology examines the sounds of *parole* not in order to learn more about the processes of speech production and perception, but to evaluate the implicit predictions that phonological representations make about phonetic behavior (cf. the discussion in Beckman and Kingston 1990). Little of what has been found is compatible with the phonetic idealizations that – as we have seen in the foregoing sections – underlie modern phonology. Indeed, there is now plenty of reason to think that there are serious problems with systematic phonetics as a theory of speech. These problems are briefly sketched here.

3.1 Systematic Phonetics as Universal Categorization

The first set of problems with systematic phonetics involves the goal of providing a universally valid taxonomy of speech sounds. Even before the advent of cheap and accessible acoustic analysis, some traditional phoneticians commented on the Eurocentric bias in the IPA's categories, but recent instrumental work makes it increasingly difficult to maintain the idea of a universal categorical taxonomy. A striking example comes from Cho and Ladefoged's careful comparative study (1999) of voice onset time (VOT) in 18 different languages. Figure 11.1 shows the mean VOT in voiceless velar stops in citation forms before non-high vowels for each of the languages; in some cases the languages in question had two such phonemes, one with short-lag ("unaspirated") and one with long-lag ("aspirated") VOT. It can be seen that there is a more or less continuous range of mean VOT values; there is certainly nothing like a cluster for unaspirated and a cluster for aspirated. The authors do suggest that the continuum might be divided up into four regions (indicated by the boxes in Figure 11.1) called "unaspirated," "slightly aspirated," "aspirated" and "highly aspirated," but this view strikes me as implausible, especially considering the relatively small size of the sample of languages. That is, it seems very likely that if we computed means from many more languages with the same methodological rigor, any apparent discontinuities in the gradual increase from one end of the VOT scale to the other would disappear.

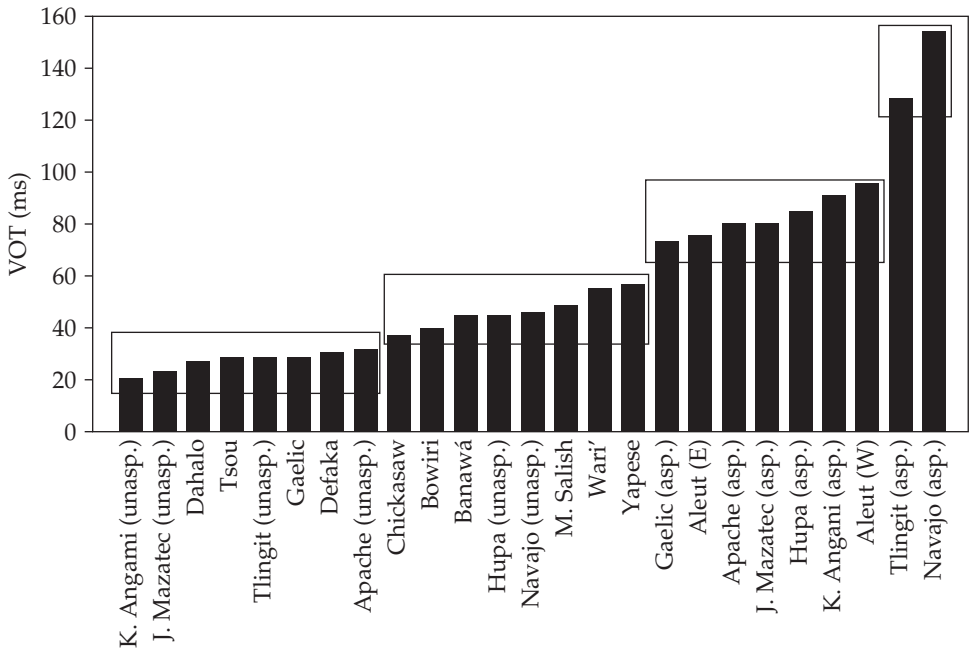


Figure 11.1 Mean voice onset time for 25 voiceless stop phonemes in 18 languages. From Cho and Ladefoged (1999).

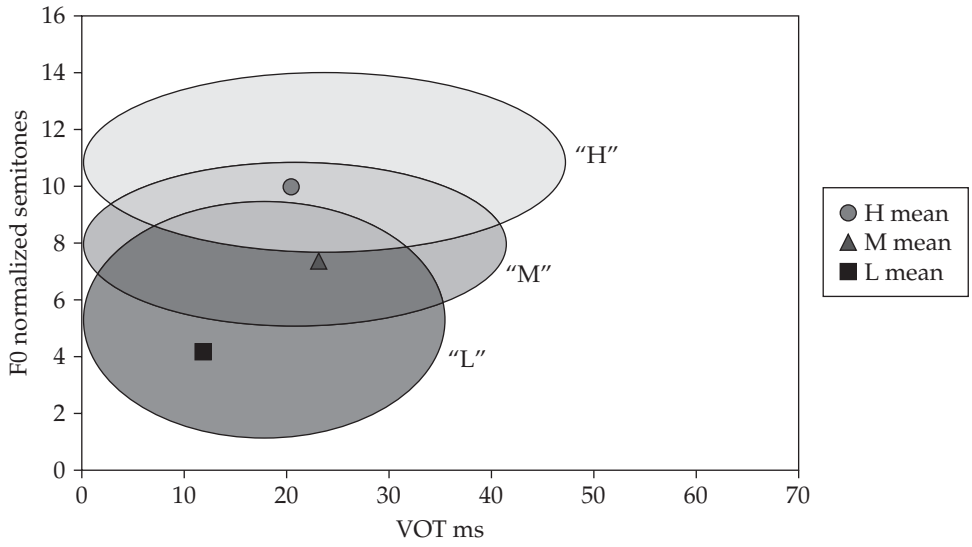


Figure 11.2 Summary plots of voice onset time and fundamental frequency for Kera syllables with high, mid, and low phonological tone. This figure is based on village women's speech; village men's speech is similar. Town-dwelling speakers influenced by French show slightly clearer VOT-based distinction between low and the other two tones. From Pearce (2007b). Reprinted with permission from Mary Pearce.

A different kind of challenge to any notion of universal categorization comes from recent work on Kera (a Chadic language spoken by some 50,000 people in Chad) by Mary Pearce (2007). According to a standard phonetic and phonological description (Ebert 1979, taken up by Odden 1994 and Rose and Walker 2004, all cited in Pearce 2007b), Kera has both voiced and voiceless stops and three distinctive tones (high, mid, and low), with various cooccurrence restrictions (in particular, voiced stops occur primarily before low tone). By analyzing the productions of several Kera speakers acoustically, however, Pearce showed that in fact VOT is extremely variable in all stops, and *co-varies with pitch*: as shown in Figure 11.2, VOT has the shortest mean in low toned syllables and is slightly longer in mid and high toned syllables, but the VOT ranges of all three tones substantially overlap. That is, VOT is not distinctive in Kera, but some of the variation in VOT is predictable from tone, and therefore, in effect, VOT is one of the phonetic cues to tone. The two-way categorization of stops as voiced or voiceless is based on the Eurocentric categories of the first phoneticians to describe the language – exactly the kind of thing Bloomfield warned against in the passage quoted earlier. Moreover, the idea that VOT could serve as a phonetic cue to the phonological category of tone cuts across a standard understanding of the distinction between segmental and suprasegmental.[6] But as Pearce amply shows, the description of the phonology of Kera makes much more sense if we adopt exactly that idea.

Both the cases just discussed could be incorporated into a modified systematic phonetic theory in which phones are defined in language-specific quantitative terms as a mean value on some measurable phonetic scale (or, more generally, as a central value in some quantitatively definable phonetic space such as the vowel space defined by the value of the first two formants). That is, we could give up the idea of *universal* categorization, but still maintain the segmental idealization and still maintain the idea that the output of the phonology is a string of systematic phones which are then passed on to physical phonetics for realization. Such realizations could be quite variable without upsetting the quantitative definition of the phone. As noted above (in the discussion of VOT allophony in English) and as just illustrated in Figure 11.2, it is now quite normal to describe the phonetic manifestation of a given phoneme in statistical terms: specifically, it is common to present such data graphically as a distribution ("cloud") of individual realization tokens in some appropriate phonetic space, and it is normal to find that the edges of such clouds overlap, even quite considerably. None of this need threaten the idea that language-specific allophones can be defined quantitatively, each with its own portion of phonetic space, as long as the overlapping distributions are statistically distinct.

However, even this idea is hard to reconcile with the results of another recent study. Flemming and Johnson (2007) investigated the acoustic realization of the two unstressed vowels found in phrases like *Rosa's roses* in American English. The two vowels are clearly distinct, in the sense that formant plots of multiple tokens of each vowel show different distributions: the second vowel of *roses* is on average higher than that of *Rosa's*, which seems to justify transcribing the two with, say, [ɨ] and [ə] respectively, as Flemming and Johnson suggest. However, the way in which the distributions overlap, shown in Figure 11.3, means that [ɨ] is essentially a subset of [ə]. There is no obvious way to reconcile this kind of distributional fact with a traditional phone-based transcription. A traditional approach might be to say that there are two distinct phones [ɨ] and [ə], one of which is used in *roses* and either of which can be used "in free variation" in *Rosa's*, and careful IPA transcriptions might represent the greater variability of *Rosa's* in exactly that way. But it can be seen that this description misrepresents the quantitative data: the distribution of the vowel in *Rosa's* appears to occupy a continuous space on the plot, not two separate spaces corresponding to two different transcriptions. That is, the quantitative data justify the statement that there are two distinct unstressed *phonemes* [ɨ] and [ə] in American English, but not that American English phonetic realizations allow us to distinguish two *phones* [ɨ] and [ə] occupying reasonably distinct areas of phonetic space.

3.2 Systematic Phonetics as Interface Representation

The second set of problems with systematic phonetics revolves around the notion of *interface*. As we have already noted, systematic phonetics is often seen, even by scholars of very different persuasions, as a *level of representation* at the *interface between the abstract and the physical*. This understanding of systematic phonetics is

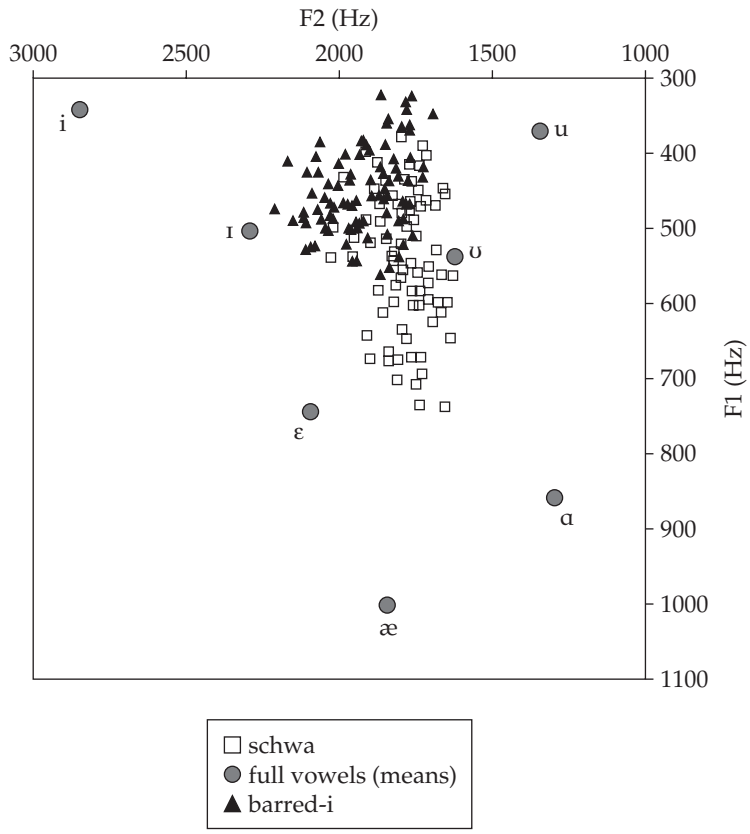


Figure 11.3 Scatterplot of formant values of individual tokens of phonemic [ə] (from Rosa's) and phonemic ɨ (from roses). Reprinted from Edward Flemming and Stephanie Johnson (2007), Rosa's roses: Reduced vowels in American English, *Journal of the International Phonetic Association*, 37, 83–96 with permission from Cambridge University Press.

made explicit in generative phonology, beginning with Chomsky (1964) and Postal (1968), but it is implicit, as Chomsky saw, in the IPA idea that there is a universally valid segmental representation of *utterances* in any language. Such an understanding is what lies behind Pike's question "Is there a significant halfway point between the continuum and the phoneme?" Some of the discussions of IPA transcription in the 1990 special issue of *Journal of Phonetics* mentioned earlier focus on its implicit claim to this interface role.

While the interface metaphor is undoubtedly somewhat misleading if taken literally as a claim about psycholinguistic processes (cf. the discussion of psycholinguistic implications of theories invoking "modularity" in Rapp and Goldrick

2000), it provides a useful way of thinking about the respective roles of *symbolic or discrete* representations and *parametric or continuous* ones in the description of language (e.g. Pierrehumbert 1990; Kornai 1994). I take it as uncontroversial that any detailed scientific description of physical processes must eventually be expressed in quantitative parametric terms. If that premise is accepted, then systematic phonetics can be interpreted as a hypothesis about the level of phonetic description *beyond which the use of symbolic representations ceases to be instructive or faithful to the empirical data*. In this light, Bloomfield's views quoted above become a competing hypothesis, namely that the level of description beyond which continuous parametric models are required is the *phonemic representation*, and that all other details of utterance phonetics cannot usefully be described in terms of symbolic categories. That is, regardless of whether the interface metaphor is ultimately enlightening psycholinguistically, there is an empirical issue here: is a symbolic idealization at the systematic phonetic level of description an appropriate part of an adequate scientific account of the sounds of language? A number of recent findings suggest that it is not.

The clearest evidence involves processes like assimilation, reduction, and neutralization. In most conceptions of phonology, these are attributed to the workings of the phonological grammar – that is, they are part of *langue*. For example, vowels before coda nasals in English are routinely said to be allophonically nasalized: one symbolic abstraction (nasal vowel) is substituted for another (oral vowel). What we actually find in the instrumental record, though, is that the nasal airflow gradually increases across the vowel, quite unlike what happens in distinctively nasal vowels in a language like French (Cohn 1993). This means that any representation in which the vowel phone is categorically represented as either nasal or non-nasal fails to express the difference between the phonetics of English and the phonetics of French. Conceivably the difference could be expressed in a systematic phonetic representation that allowed the transcriber to indicate different degrees of features like nasality, as suggested by Chomsky and Halle in *SPE* (1968: 65). However, that still precludes representing the time course of the velic opening, or any difference between the ranges of variability in the two languages. Similar comments apply to the usual conception of assimilations and deletions/reductions in connected speech, which are routinely represented as categorically either occurring or not occurring, and considered to be the output of the phonological grammar, for example, /tɛn pæst tu/ → [tɛmpæstu]. A great many studies since the mid-1980s make it clear that such representations are an extremely crude reflection of the phonetic facts (e.g. Browman and Goldstein 1986; Nolan 1992; Zsiga 1997); there are many intermediate realizations, and it seems unlikely that sharp boundaries can be established between one categorical phone-based representation and another.[7]

These cases are directly relevant to the place of systematic phonetics within *langue*. In the view made explicit in *SPE*, the phonological grammar generates a detailed (but still symbolic and segmental) phonetic representation that contains complete information about assimilations and neutralizations and the like. This detailed phonetic representation is what is passed on to the physical realization

system. The new phonetic findings suggest an alternative view: the grammar generates a rather more abstract interface representation – one that does not include any of the connected speech effects – and the interface representation is then passed on to a rather more elaborate physical realization system that specifies most aspects of pronunciation that are not the basis of categorical lexical distinctions. In such a conception of the sound system of a language, in effect, the phonology plays a smaller role in the description, while the role of phonetics is greater. The overall goal remains the same – to account for the fact that elements can count as the same in *langue* while exhibiting considerable systematic variability in their physical manifestations – but the interface between the symbolic system and the physical system is located in a different place. It seems reasonable to suggest that Bloomfield might have espoused such a view; more specifically, it seems that something like a classical phonemic transcription might serve as the "rather more abstract interface representation" that such a view requires.

None of the foregoing should be taken to suggest that the interface issue is purely a matter of efficient modeling or scientific description, devoid of psycholinguistic implications. On the contrary, the idea that the boundary between phonology and phonetics involves a representation less detailed than a systematic phonetic one is strengthened by evidence from what has been called covert contrast. First-language acquisition data is often said to involve neutralizations of adult contrasts, and various so-called phonological disorders involve children's alleged failure to distinguish adult phonemes, for example, velar and coronal stops. However, detailed instrumental investigation (e.g. Macken and Barton 1980; Scobbie *et al.* 2000) suggests that in such cases children are sometimes – perhaps usually – aware of the phonological distinction and actually produce distinct patterns of articulation which, however, are not perceived as distinct by adults (including phone-based transcribers). The contrast is thus present in the child's *phonology*, but covertly, hidden from the observer equipped only with native speaker perceptual categories. In many cases it is impossible to characterize the way the child manifests the contrast in segment-based terms, but only in terms of continuous quantitative parameters.

The case of covert contrast shows that conceiving of systematic phonetics as an interface representation has concrete consequences for our understanding of developmental disorders and for the design of appropriate therapies. If the mapping from underlying representations to phones is part of *langue*, then children's phonological disorders are appropriately named, and physical phonetics – the motor behavior involved in realizing the linguistically specified output – is irrelevant theoretically and therapeutically. However, since it appears that phonologically disordered children are actually aware of the linguistic distinction and are unable to master the appropriate motor control to produce distinguishable acoustic output, then therapy obviously needs to focus on the physical, not the linguistic. And this, once again, suggests that the level of description corresponding to a systematic phonetic representation is not the right place to locate the interface between the categorical and the continuous.

3.3 What Systematic Phonetics Could be a Theory Of

In order to avoid a potential misunderstanding, I should make clear that my remarks here are not intended as a blanket rejection of the IPA enterprise. As a tool for linguistic typology, systematic phonetics has an important role to play: terms like "front rounded vowel" and "uvular fricative" have reasonably clear language-independent definitions, and it is certainly meaningful to say that French and German have front rounded vowels while English and Spanish do not. Given what we now know about phonetic variability, statements like these must presumably be interpreted in something like the following way: French and German have phonological elements whose typical or canonical phonetic realization is a front rounded vowel, whereas English and Spanish do not, and any portion of an English or Spanish sentence that might be classed phonetically as a front rounded vowel is to be interpreted as the realization of some other phonological element. But whatever refinements of interpretation we wish to introduce into our understanding of phonetic typology, I believe that statements of this sort are useful scientific generalizations *about languages*. The problems discussed in the preceding two subsections arise from trying to use systematic phonetic terminology and concepts as descriptions of *individual acts of speech*.

The difficulty here is part of a more general problem with linguistic typology, better known from attempts to give language-independent definitions of parts of speech and of grammatical notions like *subject*. The relation between definitions that are useful for typology and those that are needed for the description of individual words and constructions in individual languages has been discussed in a number of recent articles (e.g. Huddleston and Pullum 2002: 31f; Newmeyer 2007; Haspelmath 2007), and the way forward is still far from clear. I have elsewhere discussed the same issue in connection with the transcription of intonation (Ladd 2008a: Section 3.2.2; 2008b), where the substantial current disagreements revolve in part around the nature of symbolic transcriptions. The point I wish to make here is simply that while systematic phonetics is of doubtful validity as the theoretical basis for describing utterance phonetics, it may be useful and important as a theory of phonetic typology. It is probably true, as noted by Pierrehumbert, Beckman, and Ladd (2000: 285), that "there are no two languages in which the implementation of analogous phonemes is exactly the same." It does not follow that systematic phonetic descriptions have no use anywhere in a scientific account of language; indeed, the very notion of "analogous phonemes" probably depends on such descriptions.

3.4 Segmental, Suprasegmental, Autosegmental

Although the discussion so far has focused almost exclusively on the phone, no discussion of phonetics in phonology would be complete without at least mentioning the problem of phonetic properties that fall outside the segmental idealization. Some such notion as "suprasegmental" or "prosodic" properties of

speech has been assumed at least since the beginning of the IPA. However, its theoretical basis is clearly shaky. It takes little thought to realize that the traditional set of suprasegmental features – stress, pitch and quantity – are quite distinct phonetically. It is not much of an exaggeration to say that suprasegmentals are most accurately defined as those features that are not normally (or perhaps, not easily) represented in a segmental phonetic transcription. As such, they are effectively a by-product of the phone idealization.

The problem of defining suprasegmentals is discussed by Lehiste in the introduction to her influential book entitled simply *Suprasegmentals* (1970). She concedes that the conventional denotation of the term – stress, pitch and quantity – is essentially only a list, and notes that "a definition is preferable to a list." She then briefly identifies three main approaches to providing a real definition. The three are based on: (a) the fact that suprasegmental features can apply over domains longer than a segment; (b) the supposed phonetic distinctness of suprasegmental features from the properties that define segmental phones; (c) the supposed need to define suprasegmentals phonetically in terms of a syntagmatic comparison within an utterance (Jakobson, Fant and Halle 1952: 13). All of these have some element of usefulness and all fail in important ways; a full discussion of this topic must be left for a different paper.

In the 1970s serious attempts were made to understand suprasegmental phonology, stimulated by consideration of issues that were brought to light by the formalization of phonology in *SPE*. As noted above, the *SPE* formalism treated utterances as ordered strings of segments and segments as bundles of unordered features. The only place for stress, pitch and quantity in this formalization was as features of specific segments. Liberman (1975) and Liberman and Prince (1977) proposed that stress could more insightfully be treated in terms of a hierarchical "metrical" structure, which necessitates adding some sort of bracketing or constituent structure to Chomsky and Halle's simple strings. Specifically with regard to stress, this proposal avoids some of the problems of phonetic interpretation that accompanied Chomsky and Halle's use of stress features on specific segments (see e.g. Vanderslice and Ladefoged 1972 for a typical reaction to the *SPE* analysis of stress). Much more fundamentally, the metrical proposal has led to a variety of theoretical ideas about constituent structure in phonology (e.g. Selkirk 1984b; Nespor and Vogel 1986; Pierrehumbert and Beckman 1988) whose potential has, in my opinion, only begun to be explored. What seems clear, however, is that no unified set of suprasegmental features emerges from the metrical perspective; rather, it appears that stress is a very different kind of phenomenon from tone and quantity (cf. Ladd 2008: Chapter 8).

About the same time that metrical phonology emerged as a response to the *SPE* treatment of stress, dissertations by Leben (1973a) and Goldsmith (1976b) tackled problems in the phonology of tone, leading to what came to be known by Goldsmith's term "autosegmental" phonology. Leben had demonstrated clearly that many ordinary phonological phenomena in tone languages are impossible to accommodate in any formalization of phonology that treats tone as a feature of a specific segment. Goldsmith pinpointed the problem as being what

he called the "absolute slicing hypothesis," the idea that the signal can be exhaustively segmented into elements that succeed one another in time. He proposed instead that tones are an instance of a new kind of phonological element – an "autosegment" – that can be located on a separate "tier" from other segments, and that within the separate tier a separate set of temporal ordering relations obtains.[8]

In principle, the notion of autosegment could have a purely formal phonological definition. Specifically, what is noteworthy about tones from the point of view of the *SPE* formalism is that they are *not linearly ordered* with respect to segmental phonemes. This could be seen as a purely abstract mathematical property: *SPE*-style phonological strings are totally ordered, but phonological strings that allow for autosegments are only partially ordered. However, just as Trubetzkoy depended on concrete phonetic features to describe abstract phonological oppositions, so in developing the phonological abstraction of tiers Goldsmith focused on the fact that the *phonetic realization* of tone is not synchronized in lockstep with the phonetic realization of segments. Once this phonetic fact took center stage, it became obvious that it applies to almost any aspect of phonetic realization, and theoretical discussion within autosegmental phonology rapidly moved on to a consideration of the coordination of phonetic events in real time. Any special phonological properties of tone of the sort that concerned Leben were submerged beneath the idea that tone behaves like any other feature for purposes of synchronization, and the exploration of how feature tiers are temporally coordinated was extended to cover essentially phonetic phenomena such as assimilation as well.

The problem with this development is one we have already alluded to: autosegmental representations are discrete symbolic representations, and are poorly adapted to describing physical events. More generally, the extension of autosegmental phonology to deal with issues such as assimilation illustrates again the field's repeated failure to separate – really separate – phonetics and phonology: it appears that many phonologists want their descriptions to account for the phonetic detail of utterances. Yet most are reluctant to consider the use of formalisms involving continuous mathematics and quantitative variables, and without such formalisms, it is doubtful that any theory can deal adequately with all aspects of the linguistic use of sound.

4 Where do We Go from Here?

Early twenty-first-century mainstream phonology, represented by Optimality Theory, has radically changed the form of its phonological grammar from the *SPE* codification, but continues to assume that the output of the grammar is a symbolic systematic phonetic representation, based on the phone concept. The broad continuity from the early days of the phonemic principle to the present is clear: despite the rather substantial theoretical upheavals of the early 1960s and the early 1990s, little has changed in the way most phonologists conceive of the interface between

language and speech. However, as I have argued at some length, phone-based idealizations of speech are increasingly difficult to reconcile with the findings of phonetic research. We should not be surprised, then, that in the past couple of decades there have been a number of radical responses to the growing mismatch between phonological theorizing and empirical results in phonetics.

One response is to reject "formal phonology." This is the explicit proposal of a polemical article by Port and Leary (2005), who blame generative views about language as a discrete formal system for the idea that "phonetic segments are formal symbol tokens." The diagnosis here is faulty: treating phonetic segments as symbol tokens is, as we have seen, the essence of IPA transcription and of twentieth-century phonology generally. All that early generative phonology did was to formalize widely held views about phonetics. It may be appropriate to criticize formal phonology for many things, but it is not valid to treat it as the source of the phone concept.

Another more radical response is, in effect, to reject phonology altogether. This is the upshot of some versions of what is often known as "exemplar theory" (Goldinger 1996; Coleman 2002). The strong version of exemplar theory proposes that lexical entries are directly encoded in memory on the basis of acoustic traces, thereby bypassing the need for any representation in terms of phonological categories at all. This idea has a respectable pedigree (e.g. Klatt 1979) and seems likely to form part of an eventual fully worked-out psycholinguistic understanding of how words are represented in the mental lexicon (see further Coleman 2003). However, there is experimental evidence that makes clear that some such phonological abstraction as the phoneme is needed: perceivers can rapidly update their acoustic memory *of individual phonemes*, not only of whole words (McQueen *et al.* 2006). Updatable phoneme-sized categories form part of the modified exemplar theories espoused by for example, Bybee (2001) and Pierrehumbert (2003); they accept the idea that fine phonetic detail is involved in lexical representations in some way, but they do not reject phonology altogether.

Within phonology, more or less the opposite response to exemplar theory is to argue for an outright divorce from phonetics. This case has been put most strongly in the recent literature by Hale and Reiss (e.g. 2000), who talk of "substance-free phonology." The attractiveness of this view is that it takes Trubetzkoy's radical talk of the distinction between phonology and phonetics seriously, as Trubetzkoy himself did not. While a substance-free *phonology* may be possible and even desirable, though, one problem with this proposal is that it shows no interest in accounting for *language-specific phonetic* facts. That is, Hale and Reiss may be right (and Port and Leary wrong) that it is both possible and desirable to idealize language – *langue* – as a discrete formal system. However, a complete description of any actual language will always have to include statements about the language-specific interfaces between the formal system and the physical world. This is true both in semantics and in phonetics. In that sense no complete description can ever be "substance-free."

Another different approach to the growing mismatch between theories based on systematic phonetics and the results of research in laboratory phonology is

that of so-called Articulatory Phonology (AP) (Browman and Goldstein 1986, 1989; Gafos 2002, and many others). AP represents a fundamental rethink of the interface notion and of the idea that phonological elements are symbolic abstractions: its elements are gestures, which are inherently quantitative abstractions and thus ideally suited to expressing the non-segmental aspects of phonetic realization that are increasingly coming to light. It seems quite clear that AP is correct in using a quantitative rather than a symbolic idealization *of phonetics*; time and physical space need to be modeled with continuous parameters, not categorical features or phones. Whether this also entails importing reference to actual time into our phonological abstractions, or getting rid of the segment-sized phoneme as an abstract element of the language system is less clear; one could imagine marrying the AP approach to phonetic realization with some sort of "substance-free" phonology, in which case some sort of interface representation more abstract than a systematic phonetic representation would be required (cf. Ladd 2006). In any case, AP has so far shown little interest in accounting for some of the symmetries in phonological patterning that are so central to the tradition that begins with Trubetzkoy.

It is thus doubtful that any one of these new approaches by itself indicates the true path to enlightenment in the phonetic sciences, but collectively they all suggest that a fundamental re-evaluation of the place of systematic phonetics in phonology is overdue. We have spent too much time as a field failing to resolve foundational issues and worrying about problems that simply disappear if seen in a different light. Unexamined acceptance of the twin assumptions of systematic phonetics – the segmental idealization and the universal categorization assumption – is certainly partly to blame, and recognizing systematic phonetic theory as an important source of confusion in phonology provides us with an opportunity to make real progress.

A number of things begin to make sense if we assume that there is no closed universal set of phonetic elements out of which utterances are built. For example, during the 1970s and 1980s it was widely supposed that infants are born with the ability to perceive all possible phonemic contrasts and gradually lose the ones they don't need in their language. We now know that adults preserve the ability to perceive some differences that are non-phonemic in their language (e.g. Best, McRoberts, and Goodell 2001), that children make various false steps (e.g. Pater, Stager, and Werker 2004), and more generally that children have to *learn* certain distinctions of their native language. This is exactly what we should expect if there is no universally valid categorization of phonetic segments, because without that categorization, the very concept "all possible contrasts" is incoherent.

Another theoretical conundrum that becomes suspect once we start questioning the validity of systematic phonetics is the problem of opacity in so-called chain shifts in first and second language acquisition, the famous "puddle puggle puzzle" (Smith 1973). There are many well-known cases in which language acquirers are said to replace [A] by [B] but at the same time replace [C] by [A], such as Smith's original example of /d/ → /g/ and /z/ → /d/, or the somewhat more complicated case of /s/ → [θ], /θ/ → [f], /f/ → [f] cited by Dinnsen and Barlow (1998). But

these cases are problematical only insofar as the identification of phones is accurate. The second case is a problem only if (a) [f] < /θ/ is identical to [f] < /f/ in the child's speech (which is doubtful, given the existence of covert contrast) and (b) [θ] < /s/ in the child's speech can reliably be identified with [θ] < /θ/ in adult speech (which is uncertain at best). Similarly, Smith's classic case is a problem only if [d] < /z/ in the child's speech can be identified with [d] < /d/ in adult speech. If the phonetic realizations are actually empirically distinct, the chain shift problem evaporates.

But much more central issues are at stake. The most conspicuously unresolved issue in phonology, in my view, is the debate over the classical phoneme that began in the late 1950s. The critiques by Halle (1959) and Chomsky (1964) deprived the traditional phoneme concept of its theoretical legitimacy, but it has nevertheless survived more or less intact for the intervening half-century, in practical applications (such as speech therapy, reading and literacy training, and speech technology), in linguistic fieldwork, and – revealingly – in beginning linguistics courses.[9] Schane's contention (1971) that generative phonology had only superficially done away with the phoneme has never been refuted (see further Ladd 2006).

Within the generative tradition, the problem of defining the classical phoneme manifests itself as the "abstractness controversy," first raised in such terms by Kiparsky (1968b). By comparison to classical phonemic analyses, *SPE*-style systematic phonemic representations tend to be "more abstract," in the specific sense that they differ more from the corresponding string of phones. Kiparsky pointed to various undesirable consequences of allowing unlimited abstractness in this sense, but given the *SPE* framework he was unable to find many principled reasons for avoiding it. The abstractness problem is inherent in the *SPE* architecture: "systematic phonemes" and phones are the same kind of formal object, namely bundles of features, and it is difficult to constrain a set of ordered feature-changing rules except ad hoc. From the *SPE* point of view, that is, classical phonemic theory amounted to little more than a collection of arbitrary restrictions on permissible phonological abstractions – yet there seemed to be no non-arbitrary basis for a different set of restrictions.

The controversy based on Kiparsky's statement of the problem was an active topic of debate for several years (e.g. Jackendoff 1975; Lightner 1975) but was never resolved. Instead, in the late 1970s it was merely put aside, as bright young theorists started working on other problems, in particular autosegmental and metrical phonology. Lexical Phonology in the 1980s (e.g. Halle and Mohanan 1985; Kaisse and Shaw 1985) was an attempt to deal with some of the problems Kiparsky had discussed, but in a certain sense did no more than rehabilitate something like a classical phonemic representation without resolving the question of the phoneme's theoretical legitimacy, and has accordingly failed to live up to its apparent promise. Within OT, the abstractness issue has not been explicitly discussed, presumably because of the notion that OT does not involve derivations or rules. However, the problem is still present, because the OT formalism is like the *SPE* formalism in the sense that it provides a way of establishing correspondences between one symbolic representation and another. The fact that OT has not

addressed the question of what the abstract "input" representations are for or how they should be determined does not mean that this is not a problem; the principal acknowledgment of this issue has been in the form of proposals to consider "output-output" constraints (e.g. Benua 2000; Burzio 2002). An OT based purely on output-output constraints, with due attention paid to the nature of the output representation, would represent a genuinely radical departure from *SPE*, but so far that has not happened.

The theoretical issues surrounding the phoneme and the abstractness of phonological representations have always been framed in the way they are because of the assumption that the point of the exercise is to map one symbolic abstraction onto another: phonemes onto phones, systematic phonemic onto systematic phonetic representation, OT input onto OT output. My goal in this paper has been to show that this assumption is a direct consequence of having based phonological theory on the IPA theory of systematic phonetics. If instead we start from the assumption that phonetic realization involves a mapping from symbolic phonological abstractions of some sort to a *continuous* signal describable in *quantitative* physical terms, then one of our primary theoretical tasks as phonologists must be to clarify the nature of those abstractions – in effect, to define the phoneme. Systematic phonetics almost certainly has a useful scientific role to play in an overall understanding of language. But I do not believe that we will make much progress in phonology until we stop trying to ground our theories in the systematic phonetic representation of individual acts of speech.

ACKNOWLEDGMENTS

Work on this paper was supported in part by an Individual Research Fellowship ("Simultaneous and Sequential Structure in Language") from the Leverhulme Trust. I thank Wayles Browne, Patrick Honeybone, John Joseph, Mits Ota, Geoff Pullum, Barbara Scholz, and Donca Steriade for comment, conversation, and pointers to the literature.

NOTES

1 In citing and discussing Trubetzkoy in this paper I have carefully compared the German original dating from 1939, referring to the 1958 German edition, with Cantineau's translation into French (Cantineau 1949) and with Baltaxe's translation into English (Baltaxe 1969), and the two translations are entered separately under the name of their translators in the list of references. In general I have taken direct quotes from Baltaxe's translation, but the definitions of phonetics and phonology given here are my own translations of Trubetzkoy's originals *Sprechaktlautlehre* and *Sprachgebildelautlehre*, which Baltaxe renders as "the study of sound pertaining to the act of speech" and "the study of sound pertaining to the system of language" (1969: 4). I have preferred the more concise formulation in part to facilitate repeated reference to the definitions.

Although Trubetzkoy's phrases do not use the Saussurean terms *langue* and *parole*, the context makes it clear that he meant to convey precisely the Saussurean dichotomy. Cantineau translates Trubetzkoy's definitions as *la science des sons de la parole* and *la science des sons de la langue* (1949: 3).

2 The abbreviation *IPA* is systematically ambiguous between "International Phonetic Association" and "International Phonetic Alphabet," the latter being the best-known manifestation of the former. Throughout this paper I consistently use the abbreviation only to refer to the association, and use *IPA alphabet* to refer to the alphabet. Editions of the *IPA Handbook* are referred to here in-text as for example, "IPA (1949)" or "IPA (1999)," because they have always been published as the work of the association, not of any specific author.

3 Actually, it is probably not quite accurate to say that phone abstractions are based on the personal equipment of individual phoneticians; it would be better to describe them as the collective effect of the personal equipment of a group of scholars who were all literate in alphabetic writing systems and all spoke more than one European language. Alphabetic literacy inclined them toward the segmental idealization; familiarity with several languages that used the same alphabet inclined them to identify cross-linguistic categories of sound like [b] and [p] and to focus their attention on specific phonetic details (like the difference between [p] and [p^h]) that were salient in the comparative description of the European languages.

4 The same soul-searching still goes on among self-identified phoneticians thoroughly familiar with the continuous parametric nature of speech. A particularly striking example is seen in Laver's defense of systematic phonetics (1994: Section 4.4), which comes close to acknowledging that a symbolic segmental representation cannot be reconciled with what we know from instrumental research.

5 When Trubetzkoy discusses the phonetic basis of oppositions he normally uses the German word *Eigenschaft*, which is quite abstract and is appropriately translated into English as *characteristic* or *property*; he seldom uses the word *Merkmal*, which is now the standard German technical term for the modern sense of "feature," and which more clearly conveys the idea of an actual mark of some sort. The English word *feature* is much more ambiguous: it can refer not only to abstract characteristics but also to specific objects or actual marks of some sort, especially in fixed collocations like "features of the landscape" or "distinguishing feature (of a person)." Cantineau generally translates *Eigenschaft* as *particularité* or *caractéristique* rather than *trait*, which is now the standard French technical expression for the modern sense of "feature"; when *Merkmal* occurs Cantineau generally renders it as *marque*. Baltaxe (1969: vi–vii), who prepared her translation in the late 1960s after the technical use of *feature* was well-established, deliberately avoided the term *distinctive feature* and carefully distinguishes "Trubetzkoy's theory of distinctive oppositions" from "[Jakobson's] theory of 'distinctive features'"; she generally renders *Eigenschaft* as *property* and *Merkmal* as *mark*. To the extent that one can carry out a non-electronic search of a text as long and as dense as *Principles*, it appears that the only place Cantineau uses the phrase *trait pertinent* is at the beginning of Chapter III, where Trubetzkoy (1939: 59) describes the "phonological content" of a phoneme as the *Inbegriff aller phonologisch relevanten Eigenschaften* (NB not *Merkmale*), which is translated as "all phonologically distinctive properties" by Baltaxe (1969: 66) and as "l'ensemble des traits phonologiquement pertinents" by Cantineau (1949: 68).

6 The distinction between segmental and suprasegmental is arguably another consequence of systematic phonetics, "suprasegmental" properties being merely those

that are left over when an utterance is divided into phones. This topic is discussed briefly in Section 3.4.

7 While this statement is certainly true of some connected speech processes, it is probably premature to conclude that all such processes involve gradiently variable output. A number of recent studies on a number of different languages suggest that it may be phonetically meaningful to distinguish sharply between assimilated and non-assimilated realizations in connected speech (e.g. Ellis and Hardcastle 2002; Ladd and Scobbie 2003; Kochetov and Pouplier 2009; Kainada 2009). It is not clear whether these apparently categorical effects in connected speech are related to purely phonetic "quantal" effects (Stevens 1989) or whether they reveal something important about the nature of the interface between the linguistic and the physical. Furthermore, the existence of gradiently variable connected speech effects does not preclude the possibility that such effects may become phonologized through language change. For example, Zsiga (1995) shows that the "assimilation" of /s/ to /ʃ/ before /j/ in English *confess your* is phonetically distinct from that in *confession*, suggesting that the /ʃ/ in *confession* is in some sense generated by the grammar whereas that in *confess your* is created by the workings of the physical realization system.

8 Precursors to the autosegment notion were discussed by Hockett (1955), especially Sections 26 and 3222.

9 A web search on the set of terms {*allophone, phoneme, phonology, introduction*} in March 2008 yielded more than 75,000 hits; many of the first hundred hits are handouts or syllabuses from beginning linguistics courses from major universities, including one from MIT's open courseware.

12 Corpora and Exemplars in Phonology

MIRJAM ERNESTUS AND
R. HARALD BAAYEN

1 Introduction

This chapter reviews the role of corpora in phonological research, as well as the role of exemplars in phonological theory. We begin with illustrating the importance of corpora for phonological research as a source of data. We then present an overview of speech corpora, and discuss the kinds of problems that arise when corpus data have to be transcribed and analyzed. The enormous variability in the speech signal that emerges from speech corpora, in combination with current experimental evidence, calls for more sophisticated theories of phonology than those developed in the early days of generative grammar. The importance of exemplars for accurate phonological generalization is discussed in detail, as well as the characteristics of and challenges to several types of abstractionist, exemplar, and hybrid models.

2 The Importance of Corpora for Phonology

2.1 Getting the Facts Right

Why are corpora becoming increasingly important as a data source for phonologists? One answer is that corpora help us bridge the gap between the analyst's conception of the data and the actual data themselves. Phonologists have formulated generalizations, some of which, as we know now, thanks to corpus-based research, do not do full justice to the data. Language appears to be much more complex than is generally assumed and this complexity is important for theories

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

of phonology as well as for theories of speech production and comprehension. By way of example, we discuss a number of corpus studies on assimilation, intonation, and language change.

Our first example concerns regressive voice assimilation in Dutch. There is broad consensus in the phonological literature that obstruents are obligatory voiced before /b/ and /d/ within prosodic words, including compounds (see, e.g. Booij 1995; Wetzels and Mascaró 2001). Thus, the compound *we*/t/ + /b/*oek* 'law' + 'book' is pronounced as *we*[db]*oek*. The exceptional combined presence of final devoicing, regressive voice assimilation, and progressive voice assimilation in Dutch has received considerable attention in the theoretical literature on the nature of the feature voice (privative or not) and the typology of voice (see, e.g. Lombardi 1999; Zonnebeld 2007). However, the data are much less straightforward when we consider what speakers actually produce by investigating speech corpora. Ernestus *et al.* (2006) extracted all 908 word tokens that according to the literature should show regressive voice assimilation from the subcorpus of read speech in the Spoken Dutch Corpus (Oostdijk 2000). Three phoneticians listened to the audio files and classified all of the obstruents as voiced or voiceless. Unexpectedly, only 43% of the clusters (instead of the predicted 100%) exhibited regressive voice assimilation. In 25% of the clusters progressive voice assimilation was observed, even though progressive assimilation is traditionally seen as impossible in these contexts. Thus, *wetboek* was sometimes also pronounced as *we*[tp]*oek*. Furthermore, no assimilation was observed for 20% of the data (*we*[tb]*oek*).

This is a first illustration showing that there can be a remarkable and disquietingly large gap between the received phonological wisdom and the actual data. This gap in turn questions the adequacy of phonological theories that build on the – supposedly – exceptional facts from Dutch. Of course, the corpus findings could be explained away by the assumption that the Dutch grammar only allows regressive voice assimilation, and that the observed cases of no assimilation and also those of progressive voice assimilation are due to performance factors. However, this would introduce an insurmountable gap between phonological competence and phonetic reality, and effectively render phonological theories unfalsifiable.

As a second illustration, consider regressive place assimilation in English. The traditional wisdom holds that alveolar word-final stops (/n, t, d/) often assimilate to the place of assimilation of the following labial or velar consonant. As a consequence, *gree*/n b/*oat* is often pronounced in connected speech as *gree*[m b]*oat* (Gimson 1970). A substantial amount of research in psycholinguistics has investigated the consequences of this assimilation process for the listener. Researchers have argued both in favor of and against a role of perceptual compensation for assimilation and its role in language acquisition (e.g. Gaskell 2003; Gow 2001; Mitterer and Blomert 2003).

Dilley and Pitt (2007) investigated regressive place assimilation in conversational English, using the Buckeye Corpus of Conversational Speech (Pitt *et al.* 2005). Regressive place assimilation was observed relatively infrequently, much less frequently than standard descriptions would lead one to believe: on average only

for 9% of their data. In contrast, deletion of the alveolar stop (32%), glottalization (15%), or unassimilated pronunciations (44%) were present more often. Again, we see that the phonologists' generalizations underestimated the complexity of the data. A phenomenon that is relatively easy to observe with minimal training in phonetics, assimilation of place, made it into the standard literature, even though it is infrequent in everyday speech. Phonological processes that are much more common in the same phonological environment went unnoticed until Dilley and Pitt's careful survey of what people actually say.

An example from the domain of intonation comes from Dainora (2001). Dainora studied downstep in American English on the basis of the Boston University Radio News (Ostendorf *et al.* 1995). Downstep refers to the phenomenon that during a sequence of high tones, the last tones may show a somewhat lower fundamental frequency, which is annotated with an exclamation mark (!H* versus H*) in Tones and Break Indices (Pierrehumbert 1980).

Do high and downstepped high tones represent two fundamentally different categories? If so, we would expect that the frequency distance between two successive high tones (H*H*) would be smaller than the distance between a high tone and a following downstepped high tone (H*!H*). On average, there is indeed such a difference. Dainora (2001), however, pointed out that the distribution of the two frequency distances appear to form one single normal distribution, with the distances between successive high tones forming the distribution's left half and the distances between high and downstepped high tones its right half. It is not the case that we have two disjunct normal distributions, one for the H*H* distances and one for the H*!H* distances. This suggests that we should not interpret !H* as a separate category in its own right, since it forms one natural continuum with H*. Instead, !H* is a marker of where the lower variants of H* occur.

Our final illustration concerns the study of rhoticity in New Zealand English by Hay and Sudbury (2005). In many dialects of English, postvocalic /r/ has been lost before consonants, and word-final /r/ is only pronounced before vowel-initial words (*car* versus *ca*[r] *alarm*). In addition to this linking /r/, these non-rhotic dialects may have intrusive /r/, which appears between vowel-final and vowel-initial words, as in *ma r and pa*. The phonological literature offers several accounts of the loss of rhoticity and the rise of linking and intrusive /r/. One theory holds that in a first stage postvocalic /r/ was lost, except in linking positions. Linking /r/ was subsequently interpreted as a sandhi-process, which gave way to intrusive /r/ (Vennemann 1972). Other researchers have argued that in non-rhotic dialects, linking /r/ spread to new words by reanalysis on the part of the listener, and that both linking /r/ and intrusive /r/ are underlyingly present (Harris 1994). Hay and Sudbury (2005) investigated the loss of rhoticity and the rise of linking and intrusive /r/ on the basis of a diachronic corpus of New Zealand English (Gordon *et al.* 2007). They found that the first generation of New Zealanders was still partly rhotic, in contrast to what is generally assumed. More surprisingly, some of these New Zealanders also showed intrusive /r/, which shows that the complete loss of preconsonantal /r/ was not necessary for the rise of intrusive

/r/ (in contrast to the first theory). Furthermore, the data show that intrusive /r/ and linking /r/ are clearly different phenomena, as intrusive /r/ is less frequent than linking /r/, and linking /r/ appears more often in high-frequency collocations and morphologically complex words, whereas intrusive /r/ is seldom found in these contexts.

All these studies clearly show that speech corpora are substantially broadening the empirical scope of phonological research. Corpora show that many well-established basic facts that constitute a kind of canon feeding both phonological theory and psycholinguistic theories involve substantial simplifications that do not do justice to the variability and the range of phenomena that characterize actual speech.

2.2 *Discovering New Facts*

Corpora are also becoming increasingly important as a data source for phonologists because they reveal new facts which we did not know were right there in our own languages. It is difficult to pay attention to the details of the acoustic signal, when we are listening to our own language, since in normal language use the focus of attention is on content instead of form. This is especially so when listening to casual speech. As a consequence, we know very little about the fine phonetic detail of words in fast, unscripted speech. Such details are relevant for phonological theory, however, as they constitute an intrinsic part of speakers' knowledge of their language.

Take for example the pronunciation of homophones, such as *time* and *thyme*. It is generally assumed that homophones have exactly the same pronunciation, and differ only in meaning. This view has informed the theory of speech production developed by Levelt and colleagues (Levelt 1989; Levelt *et al.* 1999). In this theory, *time* and *thyme* have separate conceptual and syntactic representations, but share the same word form representation. In this model, there is no way in which the difference in meaning between *time* and *thyme* can be reflected in speech. Yet this is exactly what Gahl (2008) observed. Gahl analyzed roughly 90,000 tokens of homophones in the Switchboard corpus of American English telephone conversations. She found that words with a high token frequency, such as *time*, tend to have shorter realizations than their low-frequency homonyms, such as *thyme*, even after having controlled for factors such as speech rate and orthographic regularity. More generally, Bell *et al.* (2003), Ayelett and Turk (2004), and Pluymaekers *et al.* (2005b) all document, on the basis of speech corpora, shorter durations of segments, syllables, and words if these linguistic units or their carriers are of a higher frequency of occurrence. Such differences in fine phonetic detail must therefore be accounted for in linguistic theories and in psycholinguistic models of speech production.

An important phenomenon that can only be well studied on the basis of speech corpora is reduced speech. Well-known by now is the phenomenon of t-deletion (e.g. Browman and Goldstein), which has been studied extensively in sociolinguistics (e.g. Guy 1980; Neu 1980). Recent research has shown, however, that reduction in everyday speech is much more pervasive than the classical example of t-deletion

would suggest. In addition to /t/, many other segments are prone to deletion, and deletion is not restricted to single segments, but may affect complete syllables. For instance, English *ordinary* is often pronounced as [ɔnri], *because* as [k^{h}z], and *hilarious* as [hlɛrɛ] (Johnson 2004). Johnson's counts, based on the Buckeye corpus, show that some form of reduction characterizes no fewer than 25% of the words in colloquial American English. An example from Dutch illustrates the wide range of possible pronunciations a word may have: *natuurlijk* 'of course' may be pronounced not only in its canonical form [natyrlək], but also as [nətyrlək], [natylək], [ntylək], [nətyk], [ntyk], [ndyk], [tylək], and [tyk], among others (Ernestus 2000). Similar observations have been made by Kohler for German (see, e.g. Kohler 1990).

These reduction processes might be argued to be phonetic variation and outside the domain of inquiry of phonology. However, what segments reduce and the extent to which they reduce seems to be subject to a variety of intrinsically phonological constraints. For instance, a high degree of reduction is observed only for words without sentential accent in utterance medial position (e.g. Pluymaekers *et al.* 2005a, b). Sometimes, reduction is made possible by prosodic restructuring (Ernestus 2000). Furthermore, although some phonotactic constraints that govern unreduced speech are relaxed for reduced speech, reduced speech nevertheless remains subject to many phonological and phonotactic constraints.

In turn, reduction provides information about phonological structure in casual speech. An interesting example is the reduction of *don't* in American English. On the basis of 135 tokens of *don't* from a corpus of conversational American English, Bybee and Scheibman (1999) showed that *don't* may be realized with schwa, but only after the word that most frequently precedes *don't*, that is, after *I*. The presence of *I* is more important than the identity of the word following *don't*, even though reduction is also more likely and greater if this following word is more frequent after *don't* (e.g. *know, think, mean*). These data suggest that there is a tighter cohesion within *I don't* than within, for instance, *don't know* or *don't think*. This is exactly the opposite of what would be expected given the syntactic structure of these phrases, which group together the two verb forms rather than the pronoun and the first verb. This corpus-based research thus supports earlier observations on possible syntax-phonology mismatches, which led to the development of Prosodic Phonology (e.g. Nespor and Vogel 1986).

As a final example of how corpora can reveal new facts, we mention the study of endangered languages. Collecting data from native speakers of minority languages without a tradition of literacy is often difficult, if not impossible. For endangered minority languages, speakers tend to be old, monolingual, and not used to carrying out tasks that require metalinguistic skills. Fortunately, story telling avoids such experimental problems, and corpora of recorded stories or dialogues may provide valuable information for the phonologist. Russell (2008) studied a corpus of oral narratives in Plains Cree. He investigated two vowel sandhi processes. He measured the formants and durations of some 450 sequences of /a#o/ that may be produced as [o:], and showed that this sandhi process is gradient and probably results from gestural overlap. The more specialized, possibly

morphosyntactically governed, coalescence of /a+i/ or /a:+i/ to [e:] (some 250 tokens), in contrast, appeared to be more categorical. Data such as these raise the theoretical question whether gradient sandhi processes are part of phonology or of phonetics.

2.3 *Understanding Phonology in its Wider Context*

The role of discourse and pragmatics in the grammar of pronunciation is becoming a more and more important field of research. An example is the study by Fox-Tree and Clark (1997). These researchers investigated the pronunciation of the definite article *the* in a corpus of spontaneous conversations. Traditional wisdom holds that the vowel of *the* is pronounced as [ə] before consonant-initial words and as [i] before vowel-initial words. Fox-Tree and Clark showed that speakers also use the realization with [i] in non-fluent speech when they are dealing with a problem in production, ranging from problems with lexical retrieval to problems with articulation. By using [i], speakers may signal that they would like to keep the floor. The same discourse effect has been observed by Local (2007) for the realization of English *so*. On the basis of a survey of tokens of *so* extracted from corpora of spontaneous speech, Local shows that this word is reduced less when speakers want to keep the floor. It is more reduced and trails off when *so* marks the end of a turn. Such subtle use of phonetic cues is part and parcel of the grammar of a native speaker of English.

Other types of pragmatic function may affect pronunciation as well. Plug (2005), for instance, discussed the Dutch word *eigenlijk* 'actually, in fact', and documented, on the basis of a corpus of spontaneous speech (Ernestus 2000), that this word is more reduced when it signals that speakers provide information which contrasts with information that they supplied previously in the discourse. If tokens of *eigenlijk* introduce information that contradicts the presuppositions attributed to the listener, they tend to be less reduced.

Corpora have also been used to study phonological variation across social groups. Keune *et al.* (2005), for instance, investigated degree of reduction in Dutch as a function of speakers' social class, gender, age, and nationality (Belgium versus the Netherlands) on the basis of the Spoken Dutch Corpus (Oostdijk 2000). The data showed a difference between men and women (with women reducing less) and differences between social classes (but only in Belgium). Furthermore, while there was on average more reduction in the Netherlands than in Flanders, degrees of reduction varied strongly with individual words. Thus, whereas *natuurlijk* 'of course' reduces more often in the Netherlands, other words with the same morphological structure, such as *waarschijnlijk* 'probably', show very similar degrees of reduction across the two countries. These differences between men and women and between Flanders and the Netherlands suggest that reduction is not just driven by articulatory processes but is in part a cultural phenomenon. Phenomena such as these raise questions about how phonological theory should account for variation in the grammars of different groups of speakers in the larger speech community.

3 Using Speech Corpora

3.1 An Overview of Speech Corpora

Speech corpora are a relatively recent data source compared to corpora of written language. Traditionally, phoneticians and phonologists based their analyses on incidental observations and carefully designed experiments. Experiments have the advantage that they offer complete control over the materials. Words, phonemes, or phrases can be placed in exactly the right contexts and can be elicited in soundproof environments, free from background noise. Experiments, however, are not without disadvantages. The amounts of data gathered tend to be small and typically cannot be re-used for different purposes. Moreover, speech styles elicited in the context of experiments tend to be formal and not spontaneous, and materials are presented in isolation or in small, often artificial, contexts. To complement experimental research, the last decades have witnessed the development of several speech corpora designed specifically for spoken (American) English and Dutch. We discuss some of the most important ones, stressing the differences in speech type and sound quality.

An important early speech corpus, the TIMIT corpus of read speech (Fisher *et al.* 1986),[1] provides the data of what can be regarded as a large production experiment. TIMIT sampled read speech (6,300 sentences) from 630 speakers from several dialect regions of the United States. Two sentences were constructed to elicit as many differences between dialects as possible. Further sentences were constructed to provide a good coverage of phone pairs. A third set of sentences was sampled from existing sources to add to the diversity of sentence types and phonetic contexts. This corpus was designed, and has been used extensively, for the development of Automatic Speech Recognition systems.

A few years later, the HCRC Map Task Corpus[2] was published (Anderson *et al.* 1992). It provides a set of 128 dialogues (18 hours of speech) that were experimentally elicited with the Map Task. In this task, the two speakers in a dialogue are provided with a map that the other cannot see. One speaker has a route marked on her map, and has to guide the other speaker so that she reproduces this route on her own map. The crucial manipulation in this experiment is that the two maps are not identical, which forces speakers to engage in extensive discussions in order to complete their task. This leads to (the repetition of) specific words (especially of the missing landmarks), corrections, questions, and so on with a high probability. For instance, by annotating a landmark picture as *vast meadow*, Anderson and colleagues targeted t-deletion. All dialogues in the HCRC Map Task Corpus are transcribed and annotated for a wide range of behaviors including gaze. Map Task corpora have also been built for many other languages, including Italian, Portuguese, Czech, Japanese, and Dutch.

In contrast to TIMIT and the HCRC Map Task Corpus, the speech sampled in the Switchboard corpus (Godfrey *et al.* 1992)[3] was under no experimental control whatsoever. This corpus comprises some 2,430 telephone conversations of on

average six minutes involving speakers who did not know each other. In all, the corpus consists of 240 hours of recorded speech with about 3 million word tokens, produced by 500 speakers, both males and females, from all major dialects of American English. The corpus is fully transcribed, and each transcript is accompanied by a time alignment file which provides estimates of the beginning and end of words. Detailed information about the speakers is also available, including age, sex, education, current residence, and places of residence during the formative years.

More recently, a corpus of spontaneous conversations has become available with a high-quality acoustic signal, the Buckeye Speech Corpus (Pitt *et al.* 2005).[4] Data were collected in a quiet room with head-mounted microphones for 40 speakers (20 men, 20 women, cross-classified by age) from Columbus, Ohio. Each speaker was interviewed for one hour, leading to a corpus of some 300,000 words. Conversations were orthographically transcribed and phonemic transcriptions were obtained with the help of automatic speech recognition software. Time stamps are available for each of the phones.

Ernestus (2000) compiled a corpus of 15 hours of conversational Dutch with 10 pairs of speakers. She selected the speakers for each pair on the criterion that they knew each other very well, in the hope that they would feel free to engage in spontaneous and lively discussion, even in a sound-proof booth, with a separate microphone for each speaker. Each recording session consisted of two parts. During the first part, the speakers talked freely about all kinds of subjects. Conversations were so free that a substantial amount of gossip was elicited. During the second part of the session, the speakers had to engage in role playing, enacting scripts in which they knew each other very well. The corpus has been transcribed orthographically, its words have been labeled for part of speech, and a broad phonemic transcription is available that has been obtained using automatic speech technology (Schupper *et al.* 2008). This corpus has been a crucial source of information for the study of reduction in spontaneous Dutch. Similar corpora have recently been compiled for French, Spanish, and Czech.[5]

An example of a recent corpus that provides speech from a wide range of spoken registers is the Spoken Dutch Corpus (Oostdijk 2000).[6] This corpus (in all some 9 million words, 800 hours of speech) includes a 2.1 million word subcorpus of spontaneous face-to-face conversations, a 900,000 word subcorpus of read speech (recorded books from the library for the blind), and a 2 million word subcorpus of telephone conversations. The spontaneous face-to-face conversations were recorded at people's homes with a single microphone, in order to optimize the likelihood of obtaining spontaneous speech. The drawback, however, is that the quality of the recordings is not optimal, due to the presence of substantial background noise. The subcorpus of read speech, however, provides very high quality sound files.

A corpus of a very different type is the ONZE corpus of New Zealand English (Gordon *et al.* 2007),[7] one of the few diachronic speech corpora. It consists of three sub-corpora: a collection of radio recordings of some 300 speakers born between 1851–1910, a collection of recordings of some 140 speakers born between 1890–1930,

and a more recent and still growing collection of recordings of speakers born between 1930–1984. All sound files come with orthographic transcriptions.

Ideally, speech corpora would be paired with video, allowing researchers to investigate the roles of gesture, gaze direction, facial expression, and so in spontaneous speech. An example of such a recent multimodal corpus is the IFA Dialog Video corpus, developed by van Son and Wesseling.[8] This corpus has recordings of maximally 15 minutes for some 50 speakers of Dutch, with orthographic transcriptions, automatically derived word and phoneme alignment, part-of-speech labeling, and annotations for gaze direction. An audiovisual corpus of read speech for English is reported by Hazen *et al.* (2004).

3.2 *Transcriptions in Speech Corpora*

A collection of speech files alone does not constitute a speech corpus. Speech corpora make the audio data accessible by means of transcriptions and links between the transcriptions and the speech files. The most basic transcription is a straightforward orthographic transcription, which serves the function of providing a search heuristic for accessing the speech files. Some corpora also provide phonological or phonetic transcriptions. Obtaining reliable phonemic or phonetic transcriptions, however, is a non-trivial enterprise.

One possible procedure is to base the transcriptions on acoustic measurements. This is an option if the features to be transcribed have obvious correlates in the acoustic signal. Most features, however, such as the voice of obstruents, are cued by different aspects of the acoustic signal (e.g. the duration of the obstruent, the duration of the preceding vowel, the presence of vocal fold vibration, and so on). When the relative contributions of the different aspects to the overall percept are not well known, and when they may vary across speakers and registers, transcriptions based on (automatically obtained) acoustic measurements are infeasible.

Transcribing utterances by ear, however, is also not a trivial task, as it requires great concentration and even then remains error prone. Moreover, human transcribers tend to be influenced by their expectations, based on the words' pronunciations in clear speech, spelling, the phonotactics of the language, and so on (e.g. Cucchiarini 1993). Vieregge (1987: 9) even argues that human transcriptions are influenced by the transcribers' expectations without exception, and are never objective reflections of reality. Along the same lines, Keating (1998) suggests that pronunciation variability is necessarily confounded with transcription variability in studies with human transcribers.

Expectations play an important role, especially when the acoustic signal is less clear, for instance due to background noise. Speech may also be less clear because speakers reduced their articulatory effort and produced smaller and overlapping articulatory gestures. In such casual speech, the reduced forms may differ substantially from their unreduced counterparts, yet transcribers will tend to hear the reduced forms as unreduced.

Since transcribing is such a difficult and subjective task, listeners often disagree about the correct transcription. Notoriously difficult is deciding on the presence

versus absence of sonorant segments (such as schwa and liquids) and about segments' voice specifications. For instance, Ernestus (2000) reported that her three transcribers disagreed about the presence versus absence of the first vowel of the word *natuurlijk* 'of course' for 58% of the 274 tokens, while they disagreed on the voicing of intervocalic plosives for 15% of the more than 2,000 cases. Similar figures have been reported by Ernestus *et al.* (2006), Coussé *et al.* (2004), and Pitt *et al.* (2005). Disagreement arises even when listeners do not provide detailed transcriptions but classify word forms roughly into predefined categories of "no to low reduction" or "high reduction" (Keune *et al.* 2005).

What to do with tokens for which transcribers disagree? One obvious solution is not to incorporate them into the analyses. If the number of problematic tokens is low, this is feasible. However, when there are many problematic cases, the number of available data points may decrease substantially, and as a consequence, the power of subsequent statistical analyses as well. Furthermore, the problematic data points may all belong to a small number of classes (e.g. high vowels, or segments preceeded by liquids, or segments in unstressed syllables) which may provide crucial information and hence should not be excluded from the analysis a priori. In fact, such data points may be of theoretical interest; for instance, they may be indicative of an ongoing sound change (Saraçlar and Khudanpur 2004).

Another way of dealing with disagreements is to ask transcribers to listen to the problematic tokens again (and again) and see whether they are willing to change their classifications. This method does not necessarily lead to more accurate transcriptions, however, since the transcribers, when listening for the second time, know each other's classifications, and the classification which is eventually accepted may not be the best one, but the one obtained from the most confident transcriber. Finally, note that even when listeners are in full agreement, this does not necessarily imply that they provide the correct transcriptions. The transcribers may all be led astray by the same expectations.

Both these procedures for handling disagreements face yet another problem, since a high degree of disagreement may indicate that the phenomenon under investigation is gradient rather than category. For instance, when studying reduction or voicing, segments can be partially deleted or partially voiced, and requesting raters to give absolute judgments may not do justice to the complexity of the data. Below, we will mention yet another way to deal with inconsistent transcriptions which is based on the use of statistics and avoids this problem in a principled way.

To what extent do automatic speech recognition (ASR) systems provide a solution for this problem? An obvious advantage is that the slow, cumbersome, and subjective work by human transcribers is replaced by a computer algorithm that will always yield the same results. Unfortunately, ASR systems need to be trained on phonetically transcribed materials and as a consequence their accuracy depends heavily on the quality of these human-made training transcriptions. Several experiments have shown that ASR transcriptions generally show a somewhat lower agreement with human transcribers than human transcribers among each other (e.g. van Bael *et al.* 2006; Wester *et al.* 2001). ASR systems have difficulties, especially

with those classifications that are notoriously difficult also for human transcribers (presence versus absence of schwa, liquids, etc.).

The field of ASR systems is still in full development. One interesting new direction is the replacement of phonemic transcriptions by continuous transcriptions of articulatory based features (e.g. King and Taylor 2000; Ten Bosch *et al.* 2006). The choice of the set of features is largely inspired by both the theory of distinctive features (Chomsky and Halle 1968) and the gestural theory of speech production (Browman and Goldstein 1992). This type of ASR systems may prove especially useful for the study of fine phonetic detail.

3.3 Analyzing Corpus Data

Corpus data should be used responsibly. Corpora are not built for looking up incidental examples, however interesting they may be. We all too easily find examples that fit the hypothesis driving our research, and we all too easily overlook examples that do not fit our theory. Moreover, it has been very well documented by now that speakers show probabilistic behavior leading to (varying degrees of) intraspeaker variation. Finding one or two tokens of a word displaying the phenomenon of interest (e.g. schwa deletion) does not provide us with information about the way the speaker normally realizes the word. These two tokens may just represent exceptional pronunciations. Furthermore, we also have to investigate where the phenomenon under study could be expected but did not occur, since our theories should account for these facts as well. As in any other domain of scientific inquiry, we have to survey all potentially relevant data.

Corpus research thus necessarily implies the inspection of very large data sets, and for this statistical analysis is indispensable. In what follows, we give a brief introduction to a technique that is of particular relevance for the analysis of corpus data: linear mixed-effects modeling (Baayen 2008; Jaeger 2008). We illustrate this general modeling tool using a small, simplified, constructed data set that mirrors part of the structure of the data set of Hay and Sudbury (2005) on postvocalic /r/ in New Zealand English that we discussed above.

Consider Table 12.1, which lists for four speakers (S1, S2, S3, S4) and for fifteen word pairs (Pair1 to Pair15) the log-transformed lexical frequency of the second word (FreqWord), the log-transformed frequency of the word pair (FreqWordPair),[9] and the number of times /r/ was observed absent and present for each of the four subjects for each word pair. Given these observations, we ask ourselves the following questions.

(i) Does the probability of the presence of /r/ decrease with FreqWord?
(ii) Does this probability increase with FreqWordPair?
(iii) Does the probability of /r/ vary between speakers?
(iv) Does the probability of /r/ vary between word pairs?

To answer these questions, we fit a regression model to the data, with as predictors, FreqWord, FreqWordPair, Speaker, and Word Pair. Our dependent

Table 12.1 Counts of non-rhotic and rhotic variants (in that order) for four subjects (S1, S2, S3, S4) for 15 word pairs (W1 . . . W15) with varying log frequencies for the second word of the pair (`FreqWord`) and for the complete word pair (`FreqWordPair`).

Word Pair	*FreqWord*	*FreqWordPair*	*S1*	*S2*	*S3*	*S4*
Pair1	4.69	0.26	0 4	0 4	0 5	0 1
Pair2	4.25	0.26	0 4	3 1	1 2	0 3
Pair3	4.21	0.45	0 4	3 1	2 2	0 6
Pair4	4.56	0.34	0 4	2 1	1 2	0 6
Pair5	4.73	0.64	0 3	7 1	4 2	0 4
Pair6	3.04	0.25	0 2	2 0	4 1	1 4
Pair7	3.26	0.46	1 1	4 0	3 1	1 7
Pair8	1.46	0.17	3 3	4 0	1 0	1 3
Pair9	4.35	0.40	2 0	5 0	7 0	2 1
Pair10	4.24	0.26	0 4	1 3	0 3	0 3
Pair11	4.00	0.21	0 3	1 3	1 4	0 2
Pair12	4.99	0.03	0 4	1 2	0 1	0 2
Pair13	3.62	0.22	0 5	1 0	2 5	0 1
Pair14	2.78	0.30	0 5	5 2	2 3	0 4
Pair15	3.91	0.54	0 5	3 1	4 3	0 4

variable requires special care. Each observation in our dataset has one of two values: present (success) or absent (failure). What we are interested in is the *probability* of an /r/ given specific values for our predictors. One possibility is to analyze the percentages of successes. Percentages (and the corresponding proportions or probabilities), however, have mathematical properties that make them unsuited for regression analysis (see, e.g. Harrell 2001; Jaeger 2008, for detailed discussion). The most important one is that percentages are bounded between 0 and 100 (and proportions and probabilities between 0 and 1). A commonly used solution is to model the logarithm of the odds ratio of the successes and failures L_{ij} for Speaker i and Word pair j:

$$(1) \quad L_{ij} = \log \frac{p}{1-p}, \; p = \log \frac{\#\text{successes}}{\#\text{failures} + \#\text{successes}}$$

The log odds ranges from minus infinity to plus infinity. When there are more successes than failures, the log odds is positive, when the number of successes is the same as the number of failures, it is zero, and when the number of successes is smaller than the number of failures, it is negative. Given a regression model for the log odds, the predictions of the model on the probability (rather than the log odds) scale can be obtained using the relation

$$(2) \quad P_{ij} = \frac{1}{1 + e^{-L_{ij}}}.$$

In what follows, we model the log odds ratio L_{ij} for Speaker i and Word Pair j as a function of baseline odds ratio β_0 (the intercept), adjustments b_i and b_j to this baseline for Speaker i and Word Pair j, together with coefficients β_1 and β_2, which represent the effects of the frequency of the second word and the frequency of the word pair. These two coefficients represent slopes, the increase in rhoticity for a unit increase in frequency.

$$(3) \quad L_{ij} = (\beta_0 + b_i + b_j) + \beta_1 \text{FreqWord}_j + \beta_2 \text{FreqWordPair}_j + \varepsilon_{ij}.$$

When we fit this linear mixed-effects model to the data in Table 12.1, we find that the slope for the frequency of the word pair is 6.6, indicating that the likelihood of rhoticity increases as this frequency increases. The slope for frequency of the second word is estimated at –1.4, which means that rhoticity is less likely as this frequency increases. The model also provides detailed information about how the likelihood of rhoticity varies from speaker to speaker, and from word pair to word pair. For instance, S4 is the least rhotic speaker of the four and Speaker S2 the most rhotic. Of the word pairs, Pair 9 is realized most often with /r/, for Pair 1 the reverse holds. Tests of significance confirm that the effects of the two frequencies are significant, and that there is significant variability between speakers and between word pairs.

Of course, the real data studied by Hay and Sudbury are much more complex, and required inclusion of predictors such as speaker's sex (men turned out to produce /r/ more often than women) and the nature of the preceding and following vowels (front vowels disfavored [r]). Such variables can be added straightforwardly to the statistical models.

Our constructed example does not do justice to the non-randomness and non-independence, in natural discourse. Pickering and Garrod (2004), for instance, call attention to various priming effects in dialogue. How a given word is actually realized often depends on how that word, or similar words, were realized in the preceding discourse. This non-independence requires special care in statistical analysis (Rietveld *et al.* 2004). In mixed-effects models, it is often possible to bring such dependencies under control with the help of longitudinal variables (De Vaan *et al.* 2007; Balling and Baayen 2008). For instance, the number of times a given word appeared with a given realization in the preceding discourse can be added as a predictor to the model.

Above, we discussed the problem that transcribing speech is a difficult and subjective task that often leads to disagreement among transcribers. Hay and Sudbury (2005) had the same analyst transcribe the same materials twice with a couple of months intervening. They included in their analysis only those cases where on both occasions the same judgment was made, and thus accepted data loss. A solution explored by Ernestus *et al.* (2006) makes use of mixed-effects modeling and considers as dependent variable the individual classifications

produced by the raters, but adds the identity of the rater as an additional factor to the model. The idea is to predict what individual listener-raters think they heard instead of aggregating over listener-raters to compute a verdict of what was actually said. The regression model determines the role of the different predictors (e.g. lexical frequency, phonological properties of the word) as well as the influence of the different listener-raters for the classifications. In other words, it is left to the regression model to handle disagreements between listener-raters.

3.4 Generalizing Data to Different Speakers

We are now in the position to address the issue of how corpus-based statistical analyses relate to the theory of grammar. One question is phrased by Newmeyer (2003: 696) as follows.

> The Switchboard Corpus explicitly encompasses conversations from a wide variety of speech communities. But how could usage facts from a speech community to which one does not belong have any relevance whatsoever to the nature of one's grammar? There is no way that one can draw conclusions about the grammar of an individual from usage facts about communities, particularly communities from which the individual receives no speech input.

Recall that the Switchboard Corpus sampled speakers from all major varieties of American English. At first sight, it does indeed seem highly implausible that data from a set of speakers of variety A would help us to understand the grammar of an individual from variety B. However, mixed-effects modeling offers us the means for carefully teasing apart what is common to all speakers and what is specific to a particular dialect. Let us return to our hypothetical data on /r/ sandhi in New Zealand English. Suppose we have not just four speakers, but 40 speakers from dialect A, 30 speakers from dialect B, and 50 speakers from dialect C. (Dialects D, E, F, . . . are not sampled.) The model that we would now fit to the data would include dialect as a third random-effect predictor modifying the intercept (b_k).

(4) $L_{ij} = (\beta_0 + b_i + b_j + b_k) + \beta_1 \text{FreqWord}_j + \beta_2 \text{FreqWordPair}_j + \varepsilon_{ijk}$.

The adjustment b_k to the intercept for Dialect k informs us about the extent to which Dialect k differs from the language as a whole. Similarly, the adjustments b_i and b_j to the intercept for speaker i word pair j give us further information about the individual differences in the rate of occurrence of postvocalic [r] for the speakers and the word pairs. The coefficients β_0, β_1, and β_2 estimated by such a model tell us what is common to all dialects and to all the different word pairs and speakers within these dialects. Crucially, information on speaker X from dialect A contributes to our estimates of these β-coefficients, and therefore to our understanding of the grammar of speaker Y from dialect B. In other words, our mixed-effects model helps us to separate out the role of dialect, the role of the individual speaker, and the role of the shared grammar.

There are many other dimensions of variation that we will need to consider in our corpus-based models. One such dimension is register, contrasting, for instance, read speech with scripted speech, telephone conversations, and face-to-face conversations. Other dimensions are time, social class, and education. There are currently no speech corpora that properly sample across all these dimensions. As a consequence, conclusions based on corpus data are by necessity conditional on the input data.

4 Abstractionist and Exemplar-based Models

Corpus-based research has made it more than obvious that pronunciation variation is inherent in natural language. We have also seen that statistical models help clarify which patterns are characteristic of a language (variant) and which are of a more idiosyncratic nature. Moreover, such models indicate which factors (sociolinguistic, phonological, morphological, etc.) help explain this variation. All this information helps the researcher to develop better linguistic and psycholinguistic models.

Broadly speaking, present-day linguistic and psycholinguistic models can be classified along a continuum with at one endpoint purely abstractionist models and at the other endpoint purely exemplar-based models. These two types of models differ in their views of the nature of linguistic generalizations and the amount of detailed knowledge that is assumed to be available in the mental lexicon.

4.1 The Nature of Linguistic Generalizations

Early generative phonology and its direct successors, including Optimality Theory (e.g. Chomsky and Halle 1968; McCarthy and Prince 1993b), are typical examples of purely abstractionist models. They assume that generalizations over the language, such as Final Devoicing and the position of word stress, are stored independently from the words in the mental lexicon in the form of abstract representations. These abstract generalizations can be applied directly to new words, such as loan words, without reference to the words from which these generalizations were previously deduced during learning. For instance, the English verb *save* is pronounced in Dutch, a language with Final Devoicing, as [sef]. According to abstractionist theories, this is due to the application of a rule of Final Devoicing that exists independently of the data. In machine learning, learning strategies that build on abstract generalization are called eager or greedy learning strategies (Daelemans and van den Bosch 2005).

Purely exemplar-based models, on the other hand, do not posit generalizations in the form of abstract rules that are stored independently from the individual words. Generalizations are extracted from the exemplars only when they are needed (see e.g. Semon 1923, the first to discuss exemplar-based models). The English verb *save* is pronounced as [sef] in Dutch because on-line checking of its nearest phonological neighbors in the Dutch lexicon ([lef] 'live', [nef] 'nephew',

[xef] 'give', . . .) reveals overwhelming and in fact exceptionless support for the /f/. Exemplar-based models are thus characterized by lazy learning: generalization is delayed until a query is made to the system. The reason for this delay is, as we shall see below, that generalization accuracy is optimal when all the exemplars ever encountered are available for consideration. Forgetting rare, low-frequency forms is harmful.

The "online checking" in exemplar-based models involves the simultaneous evaluation of all relevant exemplars in memory. This imposes a large computational burden. Two different approaches have been explored. Skousen (2002) has developed algorithms for his computationally highly demanding theory of analogical modeling of language that anticipate the advent of quantum computing. Even for computationally less demanding algorithms, measures have to be taken to speed up processing. In machine learning, it is common to use tree-based memory structures that may afford compression rates of 50% or more, and hence allow shorter searches and faster retrieval of the nearest neighbors (see, e.g. Daelemans and van den Bosch 2005: 47). To increase the speed of evaluation at run-time even more, generalizations can be built into the tree-based memory, but, as we shall see below, this tends to go hand in hand with a decrease in the quality of the generalizations derived from the model (Daelemans and van den Bosch 2005: 67–73). In short, the hybrid solution trades quality for speed. We will return to this hybrid approach below.

In what follows, the focus of our discussion will be on models assuming exemplars at some linguistic level, as purely abstractionist models are presented in detail in the other chapters of this book. Furthermore, due to limitations of space, only the main properties of different types of models are discussed. We also challenge the traditional conception of phonology as a sub-discipline of pure linguistics. Many phonologists working within abstractionist frameworks view their task as developing a theory of just the declarative knowledge one must know as a speaker of a language. We see many problems with such a conception of the field. First, it is unclear what data fall under the "jurisdiction" of the phonologist. In the preceding section, we reviewed a wide range of phenomena that illustrate subtle aspects of the knowledge that speakers have about the sound structure of their language. Some of these phenomena can be explained with the theoretical apparatus of traditional phonology, others, however, will require this field to broaden its scope. Second, science in the twenty-first century is increasingly becoming an interdisciplinary endeavor. The likelihood that phonology will make significant advances while dismissing recent achievements in other fields, be it computational linguistics, psycholinguistics and neurolinguistics, or phonetics, as irrelevant, is in our view unnecessarily small.

4.1.1 The Importance of Many Exemplars Purely abstractionist models assume that a relatively small sample of exemplars is sufficient for developing robust generalizations. In this approach, once a generalization has been established, further incoming evidence has no role to play and is disregarded. By contrast, exemplar-based models assume that generalizations are most precise when based

on as large an instance base as possible. Importantly, several studies have shown that generalizations based on all available evidence are indeed better predictors of speakers' behavior (see, e.g. Daelemans *et al.* 1999). By taking more examples into account, more specific generalizations become possible, enabling exemplar-based models not only to replicate the general regularities captured by traditional grammars, but also to formulate more local, detailed regularities. Such more restricted regularities are important because they allow us to predict for which words speakers are uncertain, and to predict forms that speakers produce even though those forms are not expected under an abstractionist account. Thus Skousen's Analogical Modeling of Language not only correctly predicts that the English indefinite article tends to be *a* before consonants and *an* before vowels, but also simulates speakers' behavior in tending to chose *a* for some vowel-initial nouns which are special due to the characteristics of the phonemes following the initial vowels (Skousen 1989).

Similarly, we have shown that the traditional description of regular past-tense formation in Dutch is too simplistic (Ernestus and Baayen 2004). It is true that most verbal stems ending in a voiceless obstruent (before the application of Final Devoicing) are affixed with [tə] and all other stems with [də], but for some verbs speakers produce non-standard forms quite often (choosing [də] instead of [tə], or vice versa). The final obstruents of these verbs have voice-specifications that are unexpected, given the other words ending in the same (type of) rhyme. For instance, the verb *dub* 'waver' is special in Dutch since it ends in a voiced bilabial stop, whereas the sequence short vowel voiceless bilabial plosive is much more frequent (e.g. in *klap*, *stop*, *nip*, *step*, *hap*). In line with this local generalization, speakers often choose *te*, instead of *de* as the past-tense allomorph. Importantly, when speakers produce standard past-tense forms for these exceptional verbs, they need more time to select the correct past-tense allomorph than when producing standard past-tense forms for non-exceptional verbs. Past-tense formation in Dutch does not only obey the high-level generalization formulated in traditional phonological models, but also more local generalizations within the words' sets of phonologically similar words.

As a final example we mention the work by Plag and colleagues on stress assignment in English compounds (Plag *et al.* 2007, 2008). Their comprehensive surveys revealed that traditional factors (such as argument structure and the semantics of the head noun) were only moderately successful in predicting the position of stress. They obtained much better predictive accuracy by considering the distribution of stress positions in the modifier and head constituent families (the sets of compounds sharing head or modifier). For instance, street names involving *street* as their right-hand member pattern alike in having leftward stress (e.g. *Oxford Street*, *Main Street*), whereas street names ending in *avenue* have rightword stress (e.g. *Fifth Avenue*, *Madison Avenue*). Similar biases for left or right stress, although often less pronounced, are found across the lexicon for other constituent families. Their conclusions harmonize well with work on the interfixes in Dutch and German compounds (Krott *et al.* 2001) and on the semantic interpretation of compounds (Gagné 2001).

Several models assuming abstract generalizations have incorporated the idea that generalizations should be based on many exemplars. Two of these have been computationally implemented: Stochastic Optimality Theory (Boersma 1998; Boersma and Hayes 2001), and Minimal Generalization Learning (Albright and Hayes 2003). Stochastic Optimality Theory, unlike most other abstractionist theories, implements a continuous learning process in which stochastic constraints are continuously updated. The Minimal Generalization Learner constructs a large set of weighted rules that are learned during training. Once learning is completed, the rules are applied on-line during "testing."

As shown by Keuleers (2008), the Minimal Generalization Learner and TIMBL are computationally equivalent, with TIMBL executing similarity-based reasoning at runtime, and the Minimal Generalization Learner executing previously learned weighted rules at runtime. This shows that at the computational level, abstractionist and exemplar-based models can be equivalent. In such cases, evaluation should be guided by how much insight and guidance the models provide given current theories across theoretical linguistics, computational linguistics, psycholinguistics, and cognitive science.

4.1.2 The Productivity of Generalizations Purely abstract models assume that all generalizations are fully productive. They are assumed to apply across the board to any input that meets their input requirements. However, several studies have argued that a generalization's productivity depends on the number of exemplars in the lexicon supporting the generalization (e.g. Bybee 2001). Regularities are in general more productive if they are supported by more exemplars. Thus, word-specific pronunciation variation, which is characterized by only a small degree of lexical support (i.e. only from the lexical item itself), tends to be unstable and to disappear in favor of variation shared with other, phonologically similar, words. Only a high frequency of occurrence can protect isolated words against regularization (e.g. Bybee 2001).

Furthermore, generalizations based on words which are more similar are more productive than generalizations based on words that are less similar. A lesser degree of similarity has to be compensated for by a greater number of exemplars (and vice versa). Thus, a single exemplar can only affect a neighboring word if the two neighbors are already highly similar (Frisch *et al.* 2001).

In contrast to models assuming abstract generalizations, exemplar-based models are able to account for the effect of the number of exemplars and the similarity among the exemplars on degree of productivity. In these models generalizations are formulated by online checking of all exemplars. Each exemplar may contribute to the generalization based on its similarity. More exemplars and exemplars showing higher similarities may lead to stronger and therefore more productive generalizations. Skousen (1989), for instance, has incorporated these mechanisms in his Analogical Modeling of Language, by distinguishing sets of exemplars which differ in their influence based on their set size, their (phonological) distance to the target word, and also the consistency among the exemplars with respect to the outcome of the generalization (e.g. voiced versus voiceless for syllable-final obstruents in Dutch).

Note that it is important to carefully distinguish between *generalization* and *abstraction* (Daelemans and van den Bosch 2005). Exemplar-based models and abstractionist models share the goal of *generalization*, of being able to predict the behavior of unseen cases, and understand how this prediction follows from past experience. The crucial difference is how this goal is achieved. In purely abstractionist approaches, individual tokens (at a given level) are used to formulate abstract rules. Once the rules have been formulated, the individual tokens considered in formulating the rules are redundant, and discarded as theoretically unimportant. By contrast, exemplar-based approaches are driven by the conviction that every token counts, and that in order to achieve maximum prediction accuracy, it is essential to carefully consider the contribution of each exemplar. Thus, perhaps the most crucial difference between abstractionist and exemplar-based models is their very different evaluation of the role of human memory in language.

4.2 The Content of the Mental Lexicon

Abstractionist models typically work with sparse lexicons, with as only exception in generative grammar the work of Jackendorff (1975). Once the linguistic generalizations of the language have been deduced from the input, the input words are no longer needed to support the generalizations. If they are morphologically complex and completely regular in all respects, they can even be removed from the lexicon, as they can always be recreated via the morphophonological generalizations. The lexicon can be as sparse as to contain only lemmas (morphologically simplex forms, such as *tree* and *school*) and morphologically complex words that are semantically, morphologically, syntactically, or phonologically irregular (e.g. *children* and *juicy*). Regular morphological derivations and inflections are always derived by means of morphophonological generalizations (see, e.g. Kiparsky 1982a; Pinker 1991).

This approach, advocated especially by generative grammar, implies that the form stored in the (mental) lexicon need not be phonotactically well-formed or identical to a form that occurs in the actual linguistic output. Take, for instance, regular plural nouns in Dutch, which consist of the noun stem and the suffix [ə] or [s]. The affixation with [ə] may lead to voice alternation of the stem-final obstruent, for instance, singular [hɔnt] *hond* 'dog' versus plural [hɔndə] *honden* 'dogs'. The [t] of [hɔnt] is predictable, since Dutch words cannot end in voiced obstruents (Final Devoicing), whereas the [d] of [hɔndə] is not (compare the plural [hɔndə] with the plural [lɔntə] 'matches'). Therefore, generative grammar is forced to assume that the stored form is /hɔnd/, from which both the singular (Final Devoicing) and the plural ([ə]-affixation) can easily be computed. This underlying form is however phonotactically illegal as a surface form (see, e.g. Booij 1981; Wetzels and Mascaró 2001).

Exemplar models differ from abstractionist models in that the lexicon is viewed as a database containing huge numbers of exemplars (see, e.g. Bybee 1985, 2001; Johnson 2004). As it is difficult, if not impossible, to determine the relevance of abstract generalizations and exemplars in the lexicon, it is not surprising that many researchers have brought evidence from language processing into the debate.

In what follows, we discuss evidence for exemplars at different linguistic levels: for regular morphologically complex words, for pronunciation variants of one and the same word, and for exemplars of individual acoustic/articulatory events.

4.2.1 Storage of Regular Morphologically Complex Words An important finding from the psycholinguistic literature is that the processing of completely regular morphologically complex words is known to be affected by their frequency of occurrence. For instance, Stemberger and MacWhinney (1988) demonstrated that speakers produce fewer errors for high frequency than for low frequency regular past-tense forms. Similarly, numerous studies have demonstrated that readers' and listeners' recognition times of regularly inflected and derived words in a wide variety of languages are affected by these forms' frequency of occurrence (e.g. Baayen *et al.* 1997; Sereno and Jongman 1997; Bertram *et al.* 2000; Baayen *et al.* 2008; Kuperman *et al.* 2009; Baayen *et al.* 2007). These form-specific frequency effects show that language users have detailed knowledge at their disposal about how likely specific forms are. Such detailed knowledge is totally unexpected from the purely abstractionist perspective, especially when abstractionist models are projected straightforwardly onto language processing (see, e.g. Pinker 1991), but harmonizes well with exemplar-based models.

Additional evidence for the storage of regular morphologically complex words comes from language change. Booij (2002) discusses the historical lengthening of short vowels (accompanied by a change in vowel quality) in Dutch open syllables. This change resulted in morphologically conditioned pronunciation variation in several noun stems. Later, the change became unproductive. If the alternation had been completely governed by an abstract generalization stored independently of the relevant nouns, the loss of the generalization should have resulted in the disappearance of all the vowel alternations governed by that generalization. This, however, is not the case: Modern Dutch still shows the alternation for some words (e.g. *sch*[ɪ]*p* – *sch*[e]*pen* 'ship' – 'ships'), words which otherwise have a fully regular plural inflection. This can only be explained if it is assumed that the different forms in a word's paradigm become entrenched in lexical memory, irrespective of whether they are regular or not (see also, e.g. Tiersma 1982).

The storage of large numbers of regular derivational and inflectional forms makes it unnecessary to posit, as in generative grammar, underlying representations that would differ from the words' actual pronunciations. If all forms of a paradigm are stored in a redundant lexicon, there is no need to assume that the stem's underlying representation contains all unpredictable properties. If both Dutch /hɔnt/ 'dog' and /hɔndə/ 'dogs' are stored in the mental lexicon, there is no need to assume that the morpheme for 'dog' is represented as /hɔnd/ with the unpredictable final /d/. Neither speakers nor listeners need to compute the plural [hɔndə] from the underlying lexical representation of *hond*, since either the plural is stored in the mental lexicon together with /hɔnt/, or the voice specification of the obstruent can straightforwardly be inferred from its nearest phonological neighbors (/vɔndə/ 'found', /mɔndə/ 'mouths', /mɑndə/ 'baskets' . . .).

4.2.2 Storage of Pronunciation Variants The wide pronunciation variation observed in speech corpora cannot be accounted for by the storage of just the canonical pronunciations of the words or word forms in the lexicon. The words stored have to be accompanied by information about their possible pronunciations. Abstractionist models assume phonological rules (or interactions of phonological constraints) which derive the possible pronunciations (during speech production) and deduce the stored representations from the observed realizations (during speech comprehension). For instance, a phonological rule of flapping specifies in which segmental (and probably social) contexts American English /t/ may be realized as a flap (e.g. in the word *butter*). Similarly a rule (possibly the same) specifies that a flap in American English maps on /t/ in lexical representations. This rule of flapping applies to hundreds of words, and therefore represents a true generalization over American English.

This account of pronunciation variation faces an important challenge. Many types of pronunciation variations are restricted to just a few words, instead of all words satisfying the structural description of the generalization, as in the case of flapping /t/. For instance, in Dutch, word-final /t/ can be absent in utterance-final position only in the word *niet* 'not', and word-final velar fricatives may be absent only in *toch* 'nevertheless' and *nog* 'still' (Ernestus, 2000). In general, we see that words are more reduced the higher their frequency of occurrence, which may lead to word-idiosyncratic pronunciation variation. In abstractionist models, word-specific pronunciations imply either word-specific rules or constraints, or the storage of several pronunciations for the same word (see, e.g. Booij 1995). A question that arises in this context is how many different words have to show the same pronunciation variation for a generalization to come into existence.

Such questions are irrelevant for models that simply assume that each word is stored in the mental lexicon together with all its possible pronunciations. The representations of all these possible pronunciations might be abstract in nature (e.g. strings of phonemes), in which case the model is close to purely abstract models. Alternatively, these representations may be abstract labels for clouds of exemplars each representing one individual acoustic/articulatory event (see below). In this case, the model is more similar to a purely exemplar-based model. In both types of models, the Dutch word *niet* is stored with the pronunciations [nit] and [ni], which "explains" why this word may occur with and without [t] in all sentence positions. Importantly, these models account for word-specific pronunciation variation as well as for the productivity of alternations displayed by a wide range of words, such as /t/ flapping in American English.

Several studies have produced experimental evidence for the storage of at least some pronunciation variants. Racine and Grosjean (2002) showed that native speakers of French can accurately estimate how often a particular word is produced with and without schwa in spontaneous speech. The correlation between subjects' estimates of the relative frequencies and the relative frequencies observed in a speech corpus was $r = 0.46$. Apparently, speakers know the likelihoods of both pronunciation variants. In a purely abstractionist approach, it might be argued that this probability information is stored with the unreduced form and affects the likelihood of the

application of a schwa deletion rule. This account implies that there must be some memory trace for the reduced form, albeit not instantiated in the form of a separate lexical representation, but in the form of a word-specific probability of schwa deletion. However, from a computational perspective, this word-specific probability is difficult to distinguish from a separate representation in an exemplar-based model.

From an exemplar-based perspective, these facts would be captured by positing that the two variants are represented by two exemplars (or two clouds of exemplars) that may have different long-term probabilities of becoming active in speech comprehension or production. Connine and colleagues (for an overview see Connine and Pinnow 2006) showed that the frequencies of pronunciation variants play a role in word recognition. Their study of the nasal flap as a pronunciation variant of /nt/ in American English showed that listeners recognize words pronounced with a nasal flap faster if these words are more often produced with a nasal flap instead of [nt] (Ranbom and Connine 2007). This illustrates once again that language users are sensitive to the probabilities of pronunciation variants.

The assumption that all pronunciation variants of a word are lexically stored is not unproblematic. In Ernestus *et al.* (2002), we showed that listeners recognize reduced word forms presented in isolation with a higher accuracy the more similar they are to the corresponding unreduced forms. Thus, we found a strong positive correlation between the number of missing sounds and the number of misidentifications ($r = 0.81$). This strongly suggests that listeners recognize reduced pronunciations, taken out of their contexts, by means of the lexical representations of the unreduced counterparts. This finding can only be explained within exemplar-based theory if we make the assumption that lexical representations are specified for the context in which they occur (see e.g. Hawkins 2003). Reduced pronunciations would then be specified as "not occurring in isolation." This specification would also explain why the number of misidentifications was much lower when the reduced pronunciations were presented in their natural contexts instead of in isolation.

4.2.3 Storage of Acoustic and Articulatory Tokens The most extreme variant of exemplar-based models assumes that the mental lexicon contains all acoustic and articulatory tokens of all words that the language user has ever encountered (e.g. Johnson 2004). The lexicon thus would contain millions of tokens of every word form, many of which hardly differ in their phonetic characteristics. The lexicon would therefore be very similar to a speech corpus itself. Tokens sharing meaning would then be organized in clouds of words (cognitive categories) and would be interconnected as in all other versions of exemplar-based theories. We will refer to this specific type of exemplar-based model as "episodic models."

Episodic models differ in another crucial characteristic from the exemplar-based models described so far. They assume that all tokens are stored with all their fine phonetic detail. In contrast, models allowing just one or a small number of lexical representations for every word, each reflecting a different pronunciation type, typically assume that lexical representations are built up from abstract symbols such as phonemes, allophones, or phonological features. Listeners are assumed

to abstract away from the details of the speech signal that cannot be captured by these abstract categories. The tacit assumption is that these details would not be relevant for higher-level generalizations. The models discussed in the previous sections are thus closer to the abstractionist model endpoint of the continuum than episodic models, which occupy the other endpoint.

Lexical representations consisting of abstract symbols, such as phonemes, are problematic because the conversion of real speech into such abstract symbols, which includes the process of speaker normalization, has proven difficult to capture. For instance, the categorization of a sound as a certain phoneme (or allophone) is determined by many factors, including segmental context, the speaker's gender, and the listener's expectations (for an overview, see e.g. Johnson 1997). Episodic models obviate the need for problematic processes such as speaker normalization by assuming that every word token is stored together with all its fine phonetic detail, including the characteristics of the speaker (e.g. high versus low voice, Northern versus Southern accent).

The assumption that human beings store all their experiences in full detail, as claimed by episodic models, is not new. It has been developed in the categorization literature, which also contains discussions of purely abstractionist (see e.g. Homa *et al.* 1979) and exemplar-based (see e.g. Nosofsky 1986) models. Exemplar-based models have been highly popular ever since the article by Medin and Schaffer (1978), but have recently been seriously criticized by Minda and Smith (2002).

The popularity of episodic models within (psycho)linguistics does not only stem from the possibility of doing without speaker normalization, but also from experimental evidence showing that listeners store token specific fine phonetic detail, including detail carrying indexical information (e.g. information about speaker identity and speech rate). For instance, Craik and Kirsner (1974) showed that words are recognized faster and more accurately when they are produced by the same voice. Likewise, Cole *et al.* (1974) found that participants are faster in determining whether two words in a sequence are identical, if these two words are recorded from the same speaker. Furthermore, Schacter and Church (1992) demonstrated that when presented with stems, participants tend to form complex words which they have heard before, especially if those complex words were produced by the same voice as the stems. For production, Goldinger (1998) reported that participants tend to mimic previously heard pronunciations in their fine phonetic detail.

One of the few episodic models that has been described in (some) detail and that can capture this experimental evidence is MINERVA, developed by Hintzman (1986), and applied to speech by Goldinger (1998). In this model, word recognition involves the activation of all phonetically similar tokens in the lexicon, proportional to their similarity to the speech input. An aggregate of all activated exemplars constitutes an echo sent to the working memory, on the basis of which the speech input is recognized. The echo contains more idiosyncratic information about the exemplars in the lexicon if there are fewer of them present, while a higher number of exemplars results in a more general echo. Repetition of (the echo of)

a low frequency word may therefore result in a token that is phonetically highly similar to one of the previously encountered tokens. Furthermore, the strength of an echo is proportional to the activation in the lexicon created by the input and a stronger echo facilitates the recognition process (and thus leads to shorter recognition times). Goldinger tested MINERVA by predicting the results of a shadowing experiment. In order to skip the first phase of the recognition process and to focus on the episodic part of the model, he converted the phonetic characteristics of the input signal and of the exemplars in the lexicon into simple vectors of numbers: Each token consisted of 100 name elements: 50 voice elements, and 50 context elements. The predictions made by MINERVA approximated the human data very closely. Thus, participants shadowed the fine phonetic detail of a stimulus more closely if they had heard only a few tokens of that word and they were faster in shadowing high frequency (compared to low frequency) words.

Another influential episodic model is XMOD, developed by Johnson (1997) for auditory word recognition. It differs from MINERVA especially in that it is an extension of the Lexical Access from Spectra (LAFS) model developed by Klatt (1979), which assumes that the incoming speech signal is transformed into a sequence of spectra (instead of vectors of abstract numbers). Johnson's XMOD assumes that during the recognition process, exemplars respond to the input in proportion to their similarity to that input. Their activation feeds activation of the abstract word nodes, which in turn enable recognition. Importantly, XMOD assumes that smaller units of linguistic structure, such as syllables and segments, emerge in the recognition process. Like word categories, these units are defined simply as sets of exemplars.

Interestingly, evidence is accumulating that when listeners make use of indexical information in previous mentions of a word, they do so only under slow processing conditions. McLennan, Luce, and Charles Luce (2005) showed this in a series of long-term repetition priming experiments, that is, lexical decision and shadowing experiments in which each target word occurred twice. Participants reacted faster on the second occurrence of a word, as expected. Importantly, the effect of identity priming was greatest if the second occurrence was similar to the first occurrence in speech rate or voice, and simultaneously also processing was slowed down, either by the nature of the nonwords in the experiment (lexical decision) or by the forced time span between the stimulus and the response (shadowing). Mattys and Liss (2008) reported similar results for an experiment in which participants listened to two series of words and had to indicate for the words in the second series whether they had heard them before. Participants were faster in identifying target words as "old" if the two occurrences were produced by the same speaker and this speaker suffered from dysarthria, which slowed down the average speed in the experiment.

4.3 Hybrid Models

All models discussed so far have either abstract representations or exemplars at a given linguistic level. In addition, various models have been developed which

assume both abstract generalizations and exemplars at the same linguistic level. We will refer to them as hybrid models. These models explicitly assume both a redundant lexicon and abstract generalizations. Several types of hybrid models have been formulated recently, but none of them have been fully implemented computationally.

One of the oldest hybrid models is the one proposed by Pierrehumbert (2002). She posits both abstract phonological representations and abstract phonological rules (e.g. prosodic final lengthening) as well as exemplar clouds associated with phonological units as exhibited in words (phonemes, phoneme sequences, and the words themselves). According to this model, speakers use all of this information during phonological encoding. Perception, in contrast, proceeds without the intervention of an abstract level, since fine phonetic detail in the speech signal, which would be abstracted away at an intermediate abstract phonological level, is known to affect the comprehension process.

McLennan *et al.* (2003) presented a hybrid model based on the Adaptive Resonance Theory (ART) of Grossberg and Stone (1986). This model assumes that an acoustic input activates chunks of lexical (words) and sublexical (allophones, features) representations. Some of the chunks are abstract (i.e. representations for words, allophones, phonological features) and others are captured by exemplars (e.g. speaker information). Chunks resonate with the input, and this resonance constitutes the listener's eventual percept. Importantly, more frequent chunks establish resonance with the input more easily and more quickly. Hence, by making the plausible assumption that more abstract representations are more frequent, McLennan and colleagues easily account for the finding that indexical information affects speech processing only when speech processing is slowed down.

McLennan and Luce (2005) already mention the possibility that the abstract representations and exemplars are stored in different parts of the brain. Goldinger (2007) discusses the Complementary Learning System (CLS) in which this is a central assumption. This model, which has been extensively developed into a computational model by O'Reilly and Rudy (2001) and Norman and O'Reilly (2003), assumes that an acoustic input first passes the cortical complex, where abstract processing takes place: the word is, among other things, divided into its segments and acquires its meaning. It then passes, with all fine phonetic detail still present, via the entorhinal cortex to the hippocampal complex, where it is matched with acoustically similar traces and is stored itself as well. The hippocampal complex is a fast-learning network, which, again via the entorhinal cortex, affects the more stable cortical complex. This cortical complex is specialized to slowly learn statistical regularities. The CLS can account for why indexical properties play a role in speech perception, especially when recognition is delayed. Listeners then react only after the acoustic input has arrived at the hippocampal complex, which processes indexical properties. Like MINERVA, the CLS does not yet have as its input realistic data: The model's input still consists of vectors with abstract numbers and letters.

The approach of Polysp (Polysystemic Speech Perception) developed by Hawkins and Smith (Hawkins and Smith 2001; Hawkins 2003) differs from the other models

in two crucial respects. First, it stresses the assumption that a memory trace not only consists of acoustic information, but also contains its multimedial context, for instance, visual information about the speaker's articulatory gestures, information about the room the speaker was in, and information about the relationship between the speaker and the listener. Second, Polysp assumes that the analysis of an acoustic input into its linguistic units (phonemes, etc.) is incidental. Circumstances dictate whether this analysis takes place at all, and if it takes place, whether the analysis precedes, coincides, or follows word recognition. Linguistic analysis may prevail, especially in adults with extensive experience of identifying formal linguistic structure, in formal listening situations. This approach can thus account for the finding that, at least under some circumstances, indexical information affects word recognition only when processing is slow. Polysp has not been computationally implemented but Hakwins provides some suggestions, including incorporation in the ART model developed by Grossberg and colleagues (e.g. Grossberg and Stone 1986). Note that this model is located on the continuum closer to the endpoint of exemplar-based models than any of the other models discussed above that assume both abstraction and exemplars.

4.3.1 Hybrid Aspects of Compressed Lexicons Current hybrid models build on the assumption that large numbers of individual exemplars are stored. Therefore, they run into the same problem that purely exemplar-based models have to face, namely, how to avoid an instance base with so many exemplars that it becomes impossible to query the instance base in real time. In actual computational memory based models, some form of data compression is implemented. The role of data compression is worthy of further theoretical discussion.

Data compression has a long history in computer science. Efficient data structures for storing words were already being discussed by Knuth in the early 1970s (Knuth 1973). Unsurprisingly, TIMBL, which is often applied to huge data sets, has implemented various compression algorithms. One of these, the "information gain tree" (IG-TREE), is especially interesting in the context of phonological generalizations with hybrid models.

An information gain tree is a kind of decision tree. Suppose we build such a decision tree in the context of predicting whether a stem-final obstruent in Dutch is voiced or voiceless in non-final syllable position. Each successive decision in the tree considers a feature (e.g. the manner of articulation of the obstruent) and splits the data according to this feature, assigning to each of its daughter nodes the most likely outcome (voiced or voiceless) given the set of exemplars governed by that node. Note that in this tree data structure, similar exemplars share similar paths down the decision tree. In an IG-TREE, the successive decisions are ordered in such a way that as we move from its root down to its leaf nodes, the decisions become less and less important (and less successful) in separating the voiced from the voiceless realizations.

Now consider how such an IG-TREE performs under different time constraints. Under severe time constraints, only a few decision nodes can be considered. As a consequence, the choice for voiced or voiceless has to be based on the most likely outcome associated with decision nodes high up in the tree. As a consequence,

this compressed memory will show rule-like behavior: the top nodes in the tree encode the highest-level generalizations. When time constraints are relaxed, more and more lower-level decisions will come into play, with at the lowest levels the individual exemplars. An exemplar memory compressed in this way has exactly the processing properties observed in the experiments of McLennan and Luce (2005) and Mattys and Liss (2008). This explanation is, however, completely different from that of the other hybrid models, which assume that abstract generalizations and exemplars are subserved by very different modules of the grammar (see also Ullman 2004). Models with data compression show that, computationally, the abstract generalizations can be part and parcel of the organization of exemplars in memory. We note here that, as mentioned above, hybrid architectures in machine learning may reduce online processing time, but may lead to somewhat degraded qualitative performance (Daelemans and van den Bosch 2005).

5 Concluding Remarks

To conclude, advances in information technology, computer science, and psycholinguistics have created new possibilities for the study of phonology. Corpus-based research and computational modeling offer exciting new tools for understanding the knowledge that speakers and listeners have of the sound structure of their language.

ACKNOWLEDGMENTS

This research was supported by a European Young Investigator award to the first author.

NOTES

1 http://www.ldc.upenn.edu/Catalog/readme_files/timit.readme.html
2 http://www.hcrc.ed.ac.uk/maptask/
3 http://www.ldc.upenn.edu/Catalog/readme_files/switchboard.readme.html
4 http://buckeyecorpus.osu.edu/php/corpusInfo.php
5 For information about these four corpora see http://mirjamernestus.ruhosting.nl/Ernestus/Corpora.html
6 http://lands.let.kun.nl/cgn/ehome.htm
7 http://www.ling.canterbury.ac.nz/onze/
8 http://www.fon.hum.uva.nl/IFA-SpokenLanguageCorpora/IFADVcorpus/
9 For frequencies, log transformations are required in order to reduce the enormous skew which is normally present in the distributions of frequencies.

13 The Place of Variation in Phonological Theory

ANDRIES W. COETZEE AND JOE PATER

1 Introduction

Over the past 15 years, the study of variation has become increasingly important in phonology. As recently as 1995, the previous edition of this *Handbook* did not have a chapter on variation. In fact, the term "variation" does not even appear in its subject index. Today, any volume that attempts to give an overview of the current status of the field of theoretical phonology cannot go without a chapter dedicated to variation. In this chapter, we review how the fortunes of variation have changed over the past 15 years, and discuss some of the issues that arise in making a place for variation in a theory of phonology.

For the purposes of this chapter, we understand the term "phonological variation" to describe a situation in which a single morpheme can be realized in more than one phonetic form in a single environment. This definition is intentionally broad. We do not take an a priori position on whether phonological variation includes only differences in terms of categories like [+/−voice], or whether it also includes sub-categorical distinctions in terms of phonetic features like voice onset time. We also do not wish to exclude from our definition of phonological variation an alternation between two forms that are too far apart to be related by a phonological derivation (i.e. variation involving a suppletive form). Our reluctance to draw thick lines around *phonological* variation is due to the well-known difficulty in identifying a principled way of separating phonology from morphology and phonetics (see chapters by Inkelas and Ladd, respectively). This difficulty is exacerbated when the details of variation are taken into account.

Since our discussion will be structured around the question of the locus of variation in phonology, we give a brief overview here of standard assumptions

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

about the architecture of this component of grammar. Within generative linguistics, the phonological component is usually at least implictly assumed to have roughly the following shape:

(1) Lexicon → Early Phonology → Late Phonology → Phonetic Implementation

We use the theory-neutral terms of early and late phonology rather than more theory-specific terms such as lexical and post-lexical phonology (Kiparsky 1982c), or the traditional division between morphophonology and phonology, which also entails a relatively specific characterization of the distinction between the levels. Within this general model, syntax supplies the morphemes through the operation of lexical insertion. Each morpheme has a single lexical form (except in cases of suppletion), which may be changed during the course of phonological derivation. The derivation begins with the application of a set of changes that we refer to as early phonology, which are then followed by a second set of changes that we refer to as late phonology. The exact content given to early and late phonology varies between different phonological theories, but typical characteristics assigned to each of them are given in (2).

(2)

	Early phonology	Late phonology
i.	Sensitive to morphology (because of direct interaction with lexicon, in which words are morphologically decomposed)	Insensitive to morphology (because this level has no contact with the lexicon)
ii.	May have exceptions (since these are encoded in the lexicon)	Exceptionless (because of lack of contact with lexicon)
iii.	Makes only categorical changes (because only categories are represented in the lexicon)	Can introduce non-categorical changes (because of its contact with phonetics, which requires richer representations)
iv.	Word bounded (because only single words can be input to this level)	Sensitive to cross-word contexts (because whole utterances are input to this level)
v.	Insensitive to factors like speech rate (because this level has no contact with phonetics)	Sensitive to factors like speech rate (because of direct contact with phonetics)

Though this modular structure with its associated typology of phonological changes is most closely associated with Lexical Phonology, nearly all work in generative phonology, including that in relatively non-derivational frameworks like OT, at least implicitly assumes much of it. We also note that theories differ in what they consider to be the domain of phonology: parts of early phonology are sometimes argued to be purely morphological or lexical (e.g. Hooper 1976; Dressler 1985; Ford and Singh 1985), and some of late phonology is often left to the phonetic component (on variation, see Hale, Kissock, and Reiss 2007). Even

though the details of this architecture, and how it is applied, vary tremendously, we adopt this broad outline as a means of structuring our discussion.

In the second section of the chapter, we discuss the view that variation is limited to late phonology. This position is made explicit in Lexical Phonology, and is in line with the phonetically gradient nature of many variable phenomena. In the third section we discuss evidence showing that variable processes have characteristics of early phonology, focusing on examples where they are conditioned by morphology. In this section, we introduce approaches to variation in OT, paying special attention to the partially ordered constraints theory of Kiparsky (1993b) and Anttila (1997a *et seq*.). In the fourth section, we discuss models of variation in OT-like theories in which constraints are placed on a numerical scale. These include Boersma's (1997, 1998) stochastic OT, as well as models of grammar in which ranking is entirely replaced by numerical weighting, as in OT's predecessor Harmonic Grammar (HG; Legendre, Miyata, and Smolensky 1990b; Smolensky and Legendre 2006). We discuss the strengths of these theories of grammar in terms of their ability to model quantitative aspects of phonological variation, and in terms of the existence and robustness of associated learning algorithms. We also discuss their relationship to the original generative theory of probabilistic grammar: the Variable Rules model of Labov (1969).

In the fifth section, we discuss evidence that the lexicon influences variation. Variable processes can apply differently to two lexical forms that are identical in terms of all relevant morphological and phonological properties. These data pose difficulties for any theory that advocates a strict separation between the lexicon and the level at which variation takes place, as well as for some OT approaches to variation and its learning. In this final section, we provide a brief account of these phenomena in terms of weighted lexically specific constraints. We also briefly discuss alternative formalizations of the influence of the lexicon on variation in HG, and ways in which this model can be extended to deal with differences between registers or styles that are often associated with variation. Because of the formal resemblance of weighted constraints models of variation to Labov's (1969) Variable Rules theory, especially to its elaboration as a model of probabilistic grammar (Cedergren and Sankoff 1974), there is reason to be optimistic about future strengthening of connections between research on phonological variation in sociolinguistics, and in generative phonological theory.

2 Variation Limited to Late Phonology

In rule-based phonology, variation is standardly handled by simply marking a rule as [+optional] (Vaux 2008). Labov (1969) suggests that this could be formalized by writing parentheses around the structural change of a rule, as in (3).

(3) Labov (1969: 737)
$X \rightarrow (B) \; / \; Y _ \; Z$

Labov goes on to propose an augmentation of this rule writing convention, so that it is possible to include contextual factors that influence the probability that the rule will apply. In his *variable rule* notation, Greek letter variables are used to indicate features that play this role (see Section 4.5 on how probability of rule application is calculated).

(4) Labov (1969: 738)

$$X \rightarrow (Y) \; / \begin{bmatrix} \alpha fea_i \\ \vdots \end{bmatrix} \begin{bmatrix} \underline{\quad} \\ \gamma fea_j \\ \vdots \end{bmatrix} \begin{bmatrix} \beta fea_k \\ \vdots \\ \delta fea_1 \end{bmatrix}$$

Although the "paradigm change" (Cedergren and Sankoff 1974) entailed by Labov's introduction of a probabilistic component to generative grammar did have a profound effect on subsequent research in sociolinguistics, it had little impact elsewhere in theoretical phonology until relatively recently. We will return to Labov's proposal, and some aspects of its subsequent development, in Section 4.5.

In Lexical Phonology (Kiparsky 1982c), it is sometimes proposed that only post-lexical rules can be subject to optional application (Kaisse and Shaw (1985: 6); Kiparsky (1985: 86); see also Donnegan and Stampe (1979: 145) for the related claim that Natural Phonology's processes, but not its rules, can be optional). Post-lexical rules show most of the characteristics that we ascribed above to late phonology. They are insensitive to morphology, exceptionless, and can be conditioned by cross-word contexts. Moreover, Kaisse and Shaw (1985: 6) connect the variability of post-lexical rules specifically to speech rate: "We also suspect that only post-lexical rules can be optional and subject to variation due to rate of speech, though this requires further investigation." As an example of the difference between lexical and post-lexical rules, Kiparsky (1985: 86) cites nasal place assimilation in English, a rule that applies both lexically and post-lexically. Kiparsky remarks that the lexical application of this rule is obligatory, while the post-lexical application is variable. Kiparsky does not mention the facts concerning assimilation at the prefix-root boundary; we briefly discuss them in Section 3.1.

(5) Nasal place assimilation in English

a. lexical = obligatory
   e[nt]*er*, **e*[mt]*er*, **e*[ŋt]*er*
   a[mb]*er*, **a*[nb]*er*, **a*[ŋb]*er*
   pra[ŋk], **pra*[nk], **pra*[mk]
b. Post-lexical = optional
   gree[n b]*ox* ~ *gree*[m b]*ox*
   i[n b]*ed* ~ *i*[m b]*ed*
   gree[n k]*ard* ~ *gree*[ŋ k]*ard*

The variable phrasal place assimilation in (5b) is also an example of a process that is sometimes claimed not to be phonological at all, but instead the result of

phonetic implementation. Barry (1985), Nolan (1992), and Ellis and Hardcastle (2002) provide articulatory evidence that for at least some speakers, the nasal in a phrase like *green box* retains to a variable degree its alveolar closure, even when it is perceived as fully labial (see Ernestus and Baayen, this volume, on other realizations of these sequences). This can be taken as evidence that the process is at least sometimes phonetic, rather than phonological, insofar as this intermediate articulation cannot be produced by a categorical phonological rule (though cf. Hayes 1995b on this example), and to the extent that this sort of gradience is taken to be diagnostic of rules of phonetic implementation, outside the domain of categorical phonology (though cf. Ohala 1990b; Flemming 2001b).

Probably the most intensively studied variable phonological process is another English example: the variable deletion of alveolar from word-final consonant clusters, which results in variation between [wɛst] and [wɛs] for a word like *west*. We return to this process again later in this chapter, and will review some of the relevant literature there. Here we want to point out that this is another process that has been claimed to be the result of phonetic implementation rather than of the variable application of a phonological rule. At first glance, *t/d*-deletion seems like the prototypical late phonological process. It is more likely to occur in casual or fast than in formal or slow speech. It is also reported to result in gradient, rather than categorical outcomes. Browman and Goldstein (1990) investigated the production of phrases like "perfect memory" – that is, where the first word ends in a [-Ct] cluster and the second words begins in a consonant. They recorded subjects reading these phrases in careful speech style and then in a more casual speech style. Simultaneously with the acoustic recording, they also collected articulatory data by X-ray. In the acoustic data they found evidence of an alveolar stop [t] in the careful but not in the casual speech condition, showing that the process is sensitive to speech style. However, the articulatory data indicated that the tongue blade moves towards the alveolar ridge to form the [t] closure in both the careful and the casual speech condition; the difference between the speech styles was in the timing of the gestures, rather than in the presence of alveolar closure. Bybee (2000: 73) uses this to argue that there is no need for a phonological account of the process: ". . . there is no variable rule of *t/d*-deletion. Rather there is a gradual process of shortening or reducing the lingual gesture." (See also Bybee 2001: 75–76.)

3 Variation in Early Phonology

We now turn to some examples of morphologically conditioned phonological variation that provide evidence that variation has characteristics of early phonology (Section 3.1). In Section 3.2, we present the OT theory of variation proposed by Kiparsky (1993b) and developed by Anttila (1997a *et seq*.), and show how it handles some of the English *t/d*-deletion data. In Section 3.3, we discuss some issues with this model, and briefly review other OT approaches to variation.

3.1 Morphologically Conditioned Variation

We introduced the variable process of *t/d*-deletion in English in Section 2 above, pointing out that it is sometimes considered to be the result of phonetic implementation rather than of variable phonological rule application, which is consistent with its categorization as a "late" process in the typology in (2). We begin this section by showing that *t/d*-deletion is conditioned by factors similar to those that condition the (apparently) categorical phonological processes typically studied by phonologists. We focus particularly on its morphological conditioning, which is sometimes claimed to be a characteristic of only early phonological processes.

Variationist sociolinguists have extensively studied *t/d*-deletion over the past three or four decades. We therefore have data on this process for many different dialects of English. In a 1989 paper, Labov synthesizes the current state of knowledge about this process. He identifies several factors that seem to influence the rate of *t/d*-deletion in every dialect that had been described up until then. Some of the factors that he identifies are given in (6). Guy (1994) and Coetzee (2004) provide updated literature reviews, which confirm Labov's original synthesis.

(6) Cross-dialectal generalizations about *t/d*-deletion in English

a. *Stress*: *t/d* is more likely to delete from an unstressed than a stressed syllable – i.e. more deletion from *sáfest* than from *resíst*.
b. *Third consonant*: more deletion from three consonant than two consonant clusters – i.e. more deletion from *whisked* than from *picked*.
c. *Preceding consonant*: the more similar the preceding consonant is to *t/d*, the more likely *t/d* is to delete – i.e. more deletion from *west* than from *left*.[1]
d. *Morphological status of t/d*: more deletion if *t/d* is part of a monomorpheme than if it functions as the past tense suffix – i.e. more deletion from *mist* than *missed*.
e. *Following segment*: the more sonorous the following segment, the less likely *t/d* is to delete – i.e. less deletion from *best work* than from *best book*.

These factors are all typical of ones that condition the application of non-variable phonological processes. A theory that provides a unified account of variable and categorical processes is thus likely to avoid considerable duplication of formal machinery (see especially Guy 1997 for discussion of this point with respect to the *t/d*-deletion data).

From the perspective of a theory that distinguishes between early and late phonology in the manner outlined in (2) above, the problematic observation is that *t/d*-deletion is conditioned by morphology. Guy (1991b) has shown that the morphological conditioning is even more fine-grained than reported by Labov (1989). Not only does *t/d* as the final consonant of a monomorpheme (*mist*, *land*) delete more frequently than *t/d* that functions as a past tense morpheme (*missed*, *banned*), but semi-weak past tense forms (*kept*, *told*) have an intermediate degree of

t/d-deletion. This pattern has been documented for many different dialects of English, and we report only a smattering of the available data in (7). Guy (1994) and Labov (2004: 15–16) provide further evidence of the robustness of this generalization.

(7) Deletion rate of *t/d* in different dialects of English

	Regular past (*missed*)	*Semi-weak past* (*kept*)	*Monomorpheme* (*mist*)
Philadelphia English (Guy 1991b)	17%	34%	38%
Chicano English (Santa Ana 1992)	26%	41%	58%
Tejano English (Bayley 1997)	24%	34%	56%

The fact that the application of the variable, gradient process of *t/d*-deletion is conditioned by morphology is counter to the strawman typology of phonological processes we laid out in (2). Guy's (1991b) analysis of this pattern is cast within a modified version of Lexical Phonology, one in which a variable rule can apply within the lexicon (even in Level 1; see Guy 1994 for discussion of other ways in which his analysis departs from classic Lexical Phonology), while Kiparsky (1993b) proposes a single-level OT analysis.

The English nasal place assimilation facts discussed by Kiparsky (1985) and in the previous section also show evidence of morphological conditioning. Assimilation is categorical within words derived with the "Level 1" prefix *in-*, as well as within underived words. Assimilation is gradient between the "Level 2" prefix *un-* and the following stem, as well as across word boundaries. Thus, the gradient process does apply within some words. One line of explanation for this and other apparent exceptions to morphological invisibility in late phonology is to invoke a prosodic difference; *un-* could be placed outside the prosodic word that contains the root, and thus be made to behave as an independent word for the purposes of assimilation. However, it is hard to see how this sort of account would generalize to the fine-grained morphological sensitivity of *t/d*-deletion.

Many more examples of variable processes that interact with morphology can be found in the literature, and some of these have been prominent in OT analyses of variation. Anttila's (1997a) much-discussed case of Finnish genitive plural allomorphy is notable in that the variation between the allomorphs shows every indication of being an early process. Not only are the alternations limited to the genitive, but as Anttila (1997a: Footnote 2) notes, it is unclear whether the changes are produced by phonological processes, or whether they are choices between stored allomorphs. Nonetheless, Anttila argues that the factors that produce preferences between variants are clearly phonological. Other well-known examples include variation between forms of reduplication in Ilokano, which was first

discussed in Hayes and Abad (1989), and later formally analyzed by Boersma and Hayes (2001) and Coetzee (2006), and the variable application of vowel harmony observed in Hungarian, in which individual stems vary in the extent to which they select harmonic and disharmonic suffixes (Hayes and Londe 2006; Hayes, Zuraw, Síptár, and Londe 2009). Given examples like these, it is clear that phonological theory is responsible for providing an account of variation; it cannot be phonetic implementation. In the next section, we begin to discuss the constraint-based analyses of variation that have recently emerged in OT and related theories.

3.2 *The Partially Ordered Constraints Theory of Phonological Variation*

In the version of OT proposed by Prince and Smolensky (1993/2004), which we will refer to as standard OT, the grammar of a language is a total ordering of a ranked set of constraints. Standard OT yields a single optimal output (Surface Representation in phonology) for each input (Underlying Representation). In independent developments, Kiparsky (1993b) and Reynolds (1994) proposed elaborations of OT that yield variation between optimal outputs over instances of evaluation. We will focus on one of Kiparsky's proposals, developed in much more detail by Anttila (1997a *et seq*.), which we will refer to as the partially ordered constraints (POC) model of variation. In this theory, the grammar states a partial, rather than a total, order on the constraint set. Each time the grammar is used to evaluate a candidate set, one of the total orders consistent with the partial order is randomly chosen. When some of these total orders pick different candidates as optimal, variation results.

In (8), we give a schematic example. In this example, the grammar is not a total ordering of the constraints – although both C_2 and C_3 are ranked beneath C_1, their relative order is unspecified. Every time that an input is submitted to the grammar, one of the possible rankings between C_2 and C_3 is chosen. As the example shows, $cand_1$ is selected as optimal under one ranking, and $cand_2$ under the other ranking. This is therefore a language where /input$_1$/ will be either $cand_1$ or $cand_2$, but will never surface as $cand_3$.

(8) Grammar: $C_1 >> C_2$, $C_1 >> C_3$

a. First possible ranking: $C_1 >> C_2 >> C_3$

/input$_1$/	C_1	C_2	C_3
☞ $cand_1$			*
$cand_2$		*!	
$cand_3$	*!		

b. Second possible ranking: C_1 >> C_3 >> C_2

/input$_1$/	C_1	C_3	C_2
cand$_1$		*!	
☞ *cand*$_2$			*
cand$_3$	*!		

The model also predicts the frequency with which different variants will be observed, according to the principle in (9), from Anttila (1997a).

(9) Quantitative interpretation of multiple rankings

Let *t* be the total number of total orders corresponding to a partially ordered constraint set. If a candidate is selected as optimal in *n* of these rankings, then this candidate's probability of occurrence is *n/t*.

In the grammar in (8), there are two possible total orders, and each of the candidates is selected as optimal under one of these rankings. Each of these candidates therefore has equal probability. In "mainstream" generative phonology, the POC model and Reynold's (1994) alternative are perhaps the first adoptions of Labov's (1969) proposal that a grammar can encode a probability distribution over outcomes (see Bod, Hay, and Jannedy 2003 for an overview of other probabilistic models of grammar, in phonology and elsewhere).

Kiparsky (1993b) proposes the POC model in the context of an analysis of the morphological and phonological factors influencing the rate of *t/d*-deletion in English (see also Reynolds 1994). Here we provide an illustration of the POC theory by providing a slightly amended version of Kiparsky's analysis of the effect of phonological context. As we showed in (6), there are many different factors that influence the probability that *t/d*-deletion will apply in a specific instance. Like Kiparsky, we focus here only on the influence of what Guy (1991a) calls the "external" context – that is, what follows the word-final *t/d*. As Labov (1989) points out, a clear generalization that emerges from the variationist literature on *t/d*-deletion is that the rate of deletion is always highest in pre-consonantal position (*west bank*). Both pre-vocalic (*west end*) and phrase-final (*west*) position often have a lower rate of deletion, with dialects varying in which of these positions most resists deletion. In (10), we give a sample of the data from the literature on the influence of the external context on *t/d*-deletion. Chicano English is an example of a dialect with more deletion in pre-vocalic than phrase-final contexts, and the other dialects all have more deletion in phrase-final than pre-vocalic position. We abstract from some aspects of the data by lumping all consonants together. Labov (1989) and Guy (1991a, 1994), amongst others, show that some consonants are

more likely than others to induce deletion. Specifically, less sonorous consonants typically are more likely to result in deletion (i.e. more deletion in *best book* than in *best week*). Syllable structure constraints may also play a role (though cf. Labov 1997). For instance, as Guy (1991a, 1997) points out, the fact that more deletion is observed before [l] than before [ɹ] (e.g. more deletion in *best luck* than in *best rock*) may be due to the fact that [tɹ-] is a possible onset cluster but [tl-] is not.[2]

(10) Percent deletion in different contexts[3]

	Pre-V *west end*	Pre-Pause *west*	Pre-C *west side*
AAVE (Washington, DC)	29	73	76
Chicano English	45	37	62
Jamaican English	63	71	85
Trinidadian English	21	31	81
Tejano English	25	46	62

Our analysis draws in particular on that of Coetzee (2004: Chapter 5). The constraints assign violation marks as follows.

(11) *Ct — Assign a violation mark to a word-final consonant cluster ending in a coronal stop.

Max — Assign a violation mark to an input consonant that is not present in the output.

Max-Pre-V — Assign a violation mark to an input consonant in pre-vocalic position that is not present in the output.

Max-Final — Assign a violation mark to an input consonant in phrase-final position that is not present in the output.

The markedness constraint (*Ct) penalizes specifically a consonant cluster that ends in a [t] or [d] (see Coetzee (2004: 252–255) for ways in which the analysis might be extended to labial and dorsal stops). Max-Pre-V and Max-Final are contextual faithfulness constraints in the spirit of Steriade's licensing by cue constraints (Steriade 2001, 2009). These constraints protect consonants from deletion where they are more perceptible, that is, in contexts where their perceptual cues are more robustly licensed. To correctly identify a consonant, it is necessary to perceive both its place and manner of articulation. Information about the place and manner of articulation of consonants is realized in the consonantal release and in the formant transitions into and out of the consonant (Lahiri *et al.* 1984; Malécot 1958; Nearey and Shammas 1987; Stevens and Keyser 1989; Sussman *et al.* 1991; Walsh and Diehl 1991, etc.). In pre-consonantal position, it is unlikely that either releases or transitions will be realized, so that there is no special faithfulness constraint that protects *t/d* from deletion in pre-consonantal position. In phrase-final position, consonant releases can be realized so that consonants are

more likely to be perceived accurately in this position and hence less likely to delete. This motivates the existence of MAX-FINAL. Before a vowel, both releases and formant transitions can be realized, motivating the existence of MAX-PRE-V.

Although both cues can be realized pre-vocalically and only one phrase-finally it does not mean that pre-vocalic position is always a better sponsor for the consonant. In pre-vocalic position, the cues can only be realized across a word boundary, while no word boundary needs to be crossed in phrase-final position. This may explain why some dialects of English have more deletion in phrase-final position and others in pre-vocalic position. Dialects with more deletion in pre-vocalic position are dialects that are less tolerant of realizing consonantal cues across word boundaries. An account that encodes perceptual factors more directly might state this as a constraint, and eliminate MAX-FINAL and MAX-PRE-V in favor of constraints on the preservation of perceptual cues themselves, but we use the MAX-FINAL and MAX-PRE-V constraints for simplicity.

In adopting a perceptually oriented analysis of *t/d*-deletion, we are following not only Coetzee's (2004) markedness-based OT account, but also Labov's (1997, 2004) hypothesis that the differences in rates of deletion across contexts are due to perceptual factors. It is likely that this account could be extended to other aspects of the process, such as the distinctions between consonants that follow *t/d* discussed by Labov (1997), but we leave that for future research.

The four constraints in (11) can give rise to five different categorical systems, given in the table in (12). The first column in this table presents the rankings that must obtain to yield a particular system. The second column gives the total number out of the 24 possible rankings that contain the crucial rankings given in the first column. The final three columns indicate whether or not deletion is observed in each of the three contexts under the crucial rankings in the first column. For example, the first line shows the situation that would hold if MAX outranks *CT. This ranking is observed in half of the 24 possible rankings, indicated by the 12 in the second column. Under this ranking, deletion is blocked in all three contexts, as indicated in the last three columns.

(12) Crucial rankings, number of corresponding total orders, outcomes

		Deletion produced?		
Crucial rankings	*Total # rankings*	*Pre-V*	*Phrase-final*	*Pre-C*
a. MAX >> *CT	12	No	No	No
b. MAX-PRE-V >> *CT >> {MAX, MAX-FINAL}	2	No	Yes	Yes
c. MAX-FINAL >> *CT >> {MAX, MAX-PRE-V}	2	Yes	No	Yes
d. {MAX-PRE-V, MAX-FINAL} >> *CT >> MAX	2	No	No	Yes
e. *CT >> {MAX, MAX-PRE-V, MAX-FINAL}	6	Yes	Yes	Yes

The generalization to be captured is that pre-consonantal position always has the highest rate of deletion, while there are cross-dialectal differences in the

relative rates of pre-vocalic and phrase-final deletion. This is reflected in the categorical patterns in (12). Every possible ranking that yields deletion in pre-vocalic and phrase-final position also yields deletion in pre-consonantal position – rows (b), (c), and (e) in the table. However, there are some rankings that yield deletion only in pre-consonantal position – row (d) (see Prince 2006 for alternative ways of presenting the implications). Kiparsky (1993b) notes that if variation is the result of speakers varying in the grammars that they use, it is impossible to have a pattern of variation in which pre-consonantal position has the lowest rate of deletion.

As an example of a POC grammar that encodes a probability distribution over outcomes according to the quantitative interpretation in (9), we can take a grammar that imposes no ranking at all on this constraint set. There is only one faithfulness constraint that protects *t/d* in pre-consonantal position (Max). Whenever Max ranks below *Ct, deletion will be observed in pre-consonantal position, so that 12/24 rankings will result in *t/d*-deletion in pre-consonantal position. Two faithfulness constraints protect *t/d* from deletion in pre-vocalic position, Max and Max-Pre-V. In this context, deletion will only be observed if both Max and Max-Pre-V rank below *Ct, so that 8/24 rankings will result in deletion in pre-vocalic position. Since there are also two faithfulness constraints that are active phrase-finally, deletion will also be observed in 8/24 rankings phrase-finally.

The predicted deletion rates for POC grammars where the ranking between some of the four constraints is fixed can be determined in a similar manner. The table in (13) gives the predictions for a sample of the possible POC grammars. The first column gives the partial ordering between the constraints that holds for each specific POC grammar. The last three columns show the number of rankings that will result in deletion in each of the three contexts, as well as the predicted probability of deletion in each context. The first POC grammar in the table is the one we discussed in the previous paragraph. The other four are examples of POC grammars with a single fixed ranking, which provide differential rates of deletion between pre-vocalic and phrase-final position. We evaluate the success of this POC analysis in accounting for the actually observed variation in different English dialects in Section 3.3 below.

(13) Probabilities of deletion from quantitative interpretation of partial orders

Partial order		*Pre-V*	*Phrase-final*	*Pre-C*
a. None	# *rankings*	8/24	8/24	12/24
	p of deletion	0.33	0.33	0.50
b. Max-Pre-V >> *CT	# *rankings*	0/12	4/12	6/12
	p of deletion	0	0.33	0.50
c. *CT >> Max-Pre-V	# *rankings*	6/12	4/12	6/12
	p of deletion	0.50	0.33	0.50

d. MAX-FIN >> *CT	# *rankings* *p of deletion*	4/12 0.33	0/12 0	6/12 0.50
e. *CT >> MAX-FIN	# *rankings* *p of deletion*	4/12 0.33	6/12 0.50	6/12 0.50

3.3 Predictions of the POC Theory

Above we pointed out that the POC theory shares with Labov's (1969) variable rules the property of defining a probability distribution over variants. Anttila (1997a), however, draws a distinction between the POC theory and variable rules models: that POC theory makes stronger predictions about the range of possible variable phonological systems. These predictions come from two sources: OT's universal constraint set, and the POC theory of probability distribution.

Because the POC model is cast within OT, it assumes a constraint set that imposes substantive limits on the range of possible phonological systems. For example, a version of the *t/d*-deletion system in which pre-consonantal deletion has a lower probability than in the other environments is ruled out by the absence of a MAX-PRE-C constraint that protects consonants in exactly that environment. This attribute is shared with all OT theories of variation that assume a universal constraint set. We further discuss this difference between OT models of variation and the Variable Rules model in Section 4.5 below.

The POC theory makes even stronger predictions than other OT models of variation about quantitative patterns, ones that appear too strong, as Boersma and Hayes (2001: 72) have pointed out. Because the quantitative interpretation of partial orders derives probabilities from the number of rankings that yield a particular pattern, the constraint set imposes restrictions not only on the relative probability of different processes, but also on the absolute probability of each of the processes themselves. For example, in the analysis of English *t/d*-deletion presented in 3.2, whether or not *t/d*-deletion occurs in pre-consonantal position is determined by the ranking of two constraints: MAX and *CT. There are therefore only three probabilities of deletion in this context that the POC theory can derive: 0, .50, and 1. One could always increase the size of the constraint set to yield other probability distributions, such as those observed in the dialects in (10), but this strategy becomes implausible very quickly. Boersma and Hayes (2001: 72) point out how it becomes particularly difficult to maintain in cases where the probability distribution between two variants is strongly skewed in favor of one of them. To model a situation where the probability distribution of the two variants is .99 vs. .01 in the POC theory, at least 100 different rankings are required, which necessitates at least five constraints (five constraints can be ranked in 5! = 120 different ways). But five unranked constraints alone would not suffice. To get the correct probability distribution, only one or two (1% of the 120) possible rankings must favor one variant, while the other variant must be favored by 118 or 119 of the possible rankings. This is a very unlikely scenario.

One possible reaction to this shortcoming of the POC approach is to remove the responsibility for producing probability distributions from the grammar. Coetzee (2004, 2006), for instance, claims that grammar only imposes relative probabilities on variants – that is, grammar dictates that one variant is more probable than another without specifying the absolute probability of the different variants. However, under such an approach, there is no grammatical difference between two systems with probability distributions between two variants of .10 and .90, and .40 and .60. In both the first variant is less probable than the second. We suspect that native speakers would express a much stronger dispreference for the less frequent variant in the first case (see Boersma and Hayes 2001, and Hayes and Londe 2006 for relevant data). Insofar as the phonological grammar is at least partially responsible for such judgments, if not also for the distributions themselves, it would be preferable to adopt a theory that can distinguish between such systems.

Another possible reaction to this weakness of the POC approach is to change the model in some way so that it can encode probability distributions beyond those allowed by POC. Section 4 is dedicated to the discussion of OT-like theories of grammar that do this, as well as to a brief discussion of their relationship to Labov's (1969) variable rules theory. The theories discussed in Section 4 place constraints on a numerical scale. We note that another approach to quantitative aspects of variation in OT is to designate a non-numerical range over which the ranking of a constraint can vary (e.g. Reynolds 1994; Hayes and MacEachern 1998; Davidson, Juscyk, and Smolensky 2004). See Boersma and Hayes (2001) for discussion of the relationship of such a theory to one that incorporates a numerical scale.

4 Probabilistic Models of Phonology with Numerically Valued Constraints

4.1 Stochastic OT

Boersma (1997, 1998) proposes an elaboration of OT that he refers to as stochastic OT. Boersma and Hayes (2001) provide an introduction to the theory and applications to several cases of phonological variation. Though we follow this tradition in calling this theory stochastic OT, we emphasize that there other versions of OT, including those discussed in the previous section, that include a stochastic component. In stochastic OT, constraints are given values along a real-numbered scale. However, each time the grammar is used to evaluate a candidate set, the values are converted to a corresponding ranking. The size of the numerical differences between the constraints is irrelevant after this conversion: if C_1 has a value greater than C_2, then the corresponding ranking is $C_1 >> C_2$, irrespective of the size of the $C_1 - C_2$ difference.

The distance between constraints on the numerical scale does play a role in the conversion process itself. Before transforming the numerical values into a ranking, each one is perturbed by adding a different positive or negative number, taken

from a normal distribution. In successive evaluations, constraints that have numerical values sufficiently close to one another will vary in their ranking. This stochastic element of the theory is called "noisy evaluation."

Stochastic OT can yield probability distributions beyond those produced by the POC theory.[4] As a simple example, we can consider the interaction of *Ct and Max. The tableaux in (14) provide numerical values for the constraints that yield probability distributions that are highly skewed in favor of deletion and retention respectively. The probabilities of the candidates were estimated by submitting the candidate set to evaluation 100,000 times for each of the two sets of constraint values, with an evaluation noise of 2.0 (using Praat's "get output distributions" function; Boersma and Weenink 2007).

(14) Skewed probability distributions in stochastic OT

	101.6	98.2
/CtC/	*Ct	Max
.10 CtC	*	
.90 CC		*

	101.6	98.2
/CtC/	Max	*Ct
.90 CtC		*
.10 CC	*	

A particularly attractive aspect of stochastic OT is the fact that it is accompanied by a learning theory, called the Gradual Learning Algorithm (GLA). The POC theory lacks a learning algorithm, which raises both theoretical and practical difficulties. On the theoretical side, it does not inherit from standard OT the attribute of possessing a provably convergent learning algorithm (see Albright and Hayes, this volume, for further discussion of these and other theories of phonological learning). The Constraint Demotion Algorithms (CDA; Tesar 1995; Tesar and Smolensky 1998; 2000) have never been extended from standard OT to the POC theory. On the practical side, an implemented learning algorithm can aid analysts in working with the theory, as does OT-Soft's implementation of the GLA and the CDA (Hayes, Tesar, and Zuraw 2003) and Praat's implementation of the GLA (Boersma and Weenink 2007). Without such help, it can be difficult to determine

whether a given set of constraints can yield an observed pattern of variation (though see Anttila and Andrus 2006 for a partial solution).

The GLA is an online error driven learner, like some versions of the CDA. The learner is presented with one correct input-output pair at a time, and it determines the optimal output for that input, given the current state of its grammar. If that generated output differs from the learning datum, learning is triggered. The GLA updates the constraint values by subtracting some value *x* from the ranking values of each constraint that is violated more in the correct form than in the learner's own "error," and adding that same value *x* to all constraints that are violated more in the error.

The constraint values in (14) were obtained by using the implementation of the GLA in Praat. In this simulation, and all of the others we report below, the constraints start out with a value of 100, and the rate of change (*x* in the last paragraph) starts out at 1. The rate of change decreases over the course of learning (by 0.1 after each of four sets of 100,000 learning trials). When provided with the learning datum /CtC/ → [CtC], a learner with this initial state, and with noisy evaluation, might parse it incorrectly, as shown in (15). This would lead to the updated set of values shown as Grammar H_1.

(15) A learning step in the GLA with stochastic OT
Grammar H_0: *Ct 100, Max 100
Values with noise: *Ct 100.4, Max 99.8
Learning Datum: /CtC/ → [CtC]
Learner's parse:

/CtC/	*Ct	Max
CtC	*!	
☞ CC		*

Grammar H_1: *Ct, Max

When provided with sufficient examples of /CtC/ → [CtC], the learner's grammar will eventually reach a state in which *Ct is far enough Max that errors become vanishingly improbable, and the constraint values cease to change. If the learner is provided with data in which the two mappings /CtC/ → [CtC] and /CtC/ → [CC] both occur, then the learner tends to converge on values that result in probability matching. That is, the learned grammar, with the same evaluation noise, will select each of the inputs with a probability matching their relative frequency in the learning data.

4.2 Noisy Harmonic Grammar

By abandoning numerical constraint weights in favor of ranking, OT distinguishes itself from its predecessor Harmonic Grammar (HG: Legendre, Miyata, and

Smolensky 1990ab, Smolensky and Legendre 2006; see Goldsmith 1990, 1993a for early phonological applications). Stochastic OT is essentially a hybrid of the two, in that it reintroduces numerical constraint values for the purposes of modeling variation and learning, while still maintaining ranked constraint evaluation. In this section, we discuss a version of this theory that uses numerical weights in evaluation as well. If numerical constraint values are required for learning and variation, it is seems natural to also use them to choose the optimal candidate. Because this theory retains stochastic OT's noisy evaluation, we refer to it as Noisy HG. Noisy HG in the form we are using it first appeared in a Praat implementation (Boersma and Weenink 2007); see Goldrick and Daland (2009) on the similar use of noise in connectionist models of speech processing, and Boersma and Pater (2008) for other references to research using Noisy HG.

In HG, the optimal candidate is the one with the highest numerical *Harmony*, which is the sum of the weighted constraint scores. For each constraint k in a constraint set of size K, the candidate's violation score (s_k) is multiplied by that constraint's weight (w_k). To yield Harmony (H), the results are then summed, indicated by the large epsilon in equation (16).

(16) A candidate's Harmony in HG

$$H = \sum_{k=1}^{K} w_k \cdot s_k$$

We adopt Legendre, Sorace, and Smolensky's (2006) convention of converting OT violation marks to negative integers and provide a tableau in which *Ct has a greater weight than Max constraint weights are indicated immediately the constraint names. Each candidate's Harmony is indicated in the rightmost column; this figure is obtained by multiplying each violation score by the weight, and then summing these. The candidate with deletion has the highest (closest to zero) Harmony, and is thus optimal. For more detailed discussion of HG and its relation to OT, including the question of whether HG overgenerates typologically relative to OT, see especially Smolensky and Legendre (2006), Pater (2009c), and Potts, Pater, Jesney, Bhatt, and Becker (2010).

(17) A weighted constraint tableau

	2	1	*H*
/CtC/	*Ct	Max	
☞ CC		–1	–1
CtC	–1		–2

Variation can be obtained just as in stochastic OT by perturbing the constraint values by noise each time the grammar is used. One way in which this theory

differs from stochastic OT is in that it is capable of producing cumulative constraint interaction. An example of cumulativity that involves variation, which is also discussed in HG terms in Pater (2009c), comes from the phonology of Japanese loanwords (Nishimura 2003; Kawahara 2006). In native Japanese words, obstruents are categorically banned (by "Lyman's Law"), which Itô and Mester (1986) account for in terms of an OCP-Voice constraint (cf. Itô and Mester 2003a). Voiced obstruent geminates are also absent in native words, motivating a *Voiced-Geminate constraint (*Vce-Gem). In loanwords, however, voiced geminates occur (e.g. [webbu] 'web') as do multiple voiced obstruents (e.g. [bobu] 'Bob'). In HG, the loanword pattern requires a weighting of the faithfulness constraint Ident-Voice above both of the markedness constraints, as shown in the pair of tableaux in (18).

(18)

	1.5	1	*H*
/bobu/	Id-Vce	OCP-Vce	
☞ bobu		–1	–1
bopu	–1		–1.5

	1.5	1	*H*
/webbu/	Id-Vce	*Vce-Gem	
☞ webbu		–1	–1
weppu	–1		–1.5

Cumulativity becomes evident in words that contain both a voiced geminate and another voiced obstruent. As Nishimura and Kawahara show, such words are subject to a process of optional geminate devoicing (e.g. [gutto] ~ [guddo] 'good') that does not affect geminates outside the Lyman's Law context. The geminate devoicing outcome is shown in (19). In this tableau, the *sum* of the violations of the constraints with lower weight, OCP-Vce and *Vce-Gem, is greater than that of the constraint with the higher weight, Id-Vce. No OT ranking of these constraints will produce this result.

(19)

	1.5	1	1	*H*
/guddo/	Id-Vce	OCP-Vce	*Vce-Gem	
guddo		–1	–1	–2
☞ gutto	–1			–1.5

By introducing noise into the evaluation process, we can model the observed variation between the outcomes in (19). The equation to calculate Harmony for a given candidate in Noisy HG is given in (20). It differs from the equation in (16) in that for each constraint, we first sample a random number from a normal distribution with mean zero (N_k), and add it to that constraint's weight before multiplying the result by the violation score.

(20) A candidate's Harmony in Noisy HG

$$H = \sum_{k=1}^{K} (w_k + N_k) \cdot s_k$$

As with stochastic OT, a set of constraint values can be obtained by submitting a data distribution to a learning algorithm, and as with stochastic OT, the currently available algorithm for Noisy HG is an on-line error-driven one, termed HG-GLA by Boersma and Pater (2008). The update rule used in the HG-GLA is broadly applied in statistical and connectionist learning frameworks, and was perhaps first used in generative linguistics by Jäger (2007) for learning of Maximum Entropy grammar (see the next section on this alternative stochastic version of HG; see Boersma and Pater 2008 for further references on the HG-GLA). The sole difference from the GLA as described above is that the degree of change for a constraint's value depends on the degree of difference between the correct form and the error. For each constraint, the difference between the number of violations in the error and in the correct form is calculated, and that difference is multiplied by a constant, and added to the constraint's weight to get the updated value. Jäger (2007) and Pater (2008) note that when the error and the correct form differ by a maximum of one violation, the update rule is identical to that of the OT-GLA.

A difference between the HG-GLA and the OT-GLA is that the HG-GLA has proofs of convergence. Fischer (2005) provides an adaptation of a stochastic gradient ascent proof for learning Maximum Entropy grammars (see the next section), while Boersma and Pater (2008) provide an adaptation of the Perceptron convergence proof, though this proof is limited to the learning of cases without variation. The stochastic OT-GLA combination has been shown to fail on a relatively simple abstract categorical pattern (Pater 2008).

We supplied a distribution of 50% devoicing for /guddo/, and consistent faithful realization for each of /bobu/ and /webbu/ to two learners implemented in Praat (we assume an even distribution between the outcomes for illustration; we have no information on their relative frequency). The first learner operated with a stochastic OT grammar, and the OT-GLA learning algorithm. The results are shown in the row labeled "St-OT" in (21). Because stochastic OT cannot represent this pattern of variation, the OT-GLA failed to converge on a set of values for the constraints. The high values shown in the columns headed by the constraint names are indicative of this non-convergence. The last three columns show the frequency of devoicing that this grammar produces for each input. This set of values does display a limited "cumulative" effect, shown in the higher frequency of devoicing for /guddo/, which will devoice if either OCP-Voice or *Vce-Gem outranks Ident-Voice. However, stochastic OT cannot produce the categorical cumulativity observed in the Japanese data, in which full devoicing of a geminate only occurs in the presence of a second voiced obstruent.[5] The second learner operated with a Noisy HG grammar, and the HG-GLA. The weighting values in the final state grammar are as described above: the sum of OCP-Voice and *Vce-Gem

equals that of IDENT-VOICE. This grammar produces a distribution that matches the frequency distribution in the learning data, as shown in (21).

(21) Grammars learned by stochastic OT and noisy Harmonic Grammar

	Frequency of devoicing in learning data			*0.0*	*0.0*	*0.50*
	OCP-VOICE	*VCE-GEM	IDENT-VOICE	bobu	webbu	guddo
St-OT	3113.9	3113.9	3113.7	0.15	0.15	0.25
N-HG	66.8	67.6	134.4	0.0	0.0	0.50

As well as demonstrating cumulative constraint interaction, the Japanese loanword example also provides a striking further demonstration that variation is sensitive to "early" phonology. The Lyman's Law restriction against multiple voiced obstruents has all the characteristics of early phonology: it is morphologically restricted (Itô and Mester 1986), and as the loanwords show, has exceptions. Nonetheless, a Lyman's Law violation contributes to the possibility of variable devoicing.

4.3 Maximum Entropy grammar

Johnson (2002) shows how an OT grammar can be transformed into a log-linear probabilistic model usually referred to as Maximum Entropy grammar; we will use the abbreviation Max-Ent-HG to emphasize that it is formally a stochastic version of HG. Goldwater and Johnson (2003), Wilson (2006a), Jäger (2007), Hayes, Zuraw, Síptár, and Londe (2009) and others apply the resulting model of variation to phonology.[6] In this section, we show how this theory relates to Noisy HG.

Max-Ent-HG calculates a probability distribution over the candidate set: the probability of a candidate is proportional to the exponential of its Harmony score. The tableau in (22) illustrates Maximum Entropy grammar with the simple case of variation in Japanese loanwords (see the above-cited papers for more complete formal presentations). Columns are added showing the result of raising e to the power of H (the exponential function), and the probability p resulting from dividing a candidate's e^H score by the sum of these scores over the candidate set.

(22)

	2	1	1	H	e^H	p
/guddo/	ID-VCE	OCP-VCE	*VCE-GEM			
guddo		–1	–1	–2	0.14	.50
gutto	–1			–2	0.14	.50

Because they both use weighted constraints, Noisy HG and Max-Ent-HG can both represent cumulative constraint interaction, which distinguishes them from stochastic OT, and standard OT. Noisy HG and Max-Ent-HG differ in that Noisy HG produces a single optimal output each time the grammar is used (modulo ties), while Max-Ent-HG defines a probability distribution over the candidate set, which is then sampled to yield an outcome for a given utterance. One result of this difference is that Max-Ent-HG gives a portion of the probability mass to a harmonically bounded candidate (Jäger and Rosenbach 2006); this cannot happen in Noisy HG if the sum of the weight and noise terms is positive (Jesney 2007). The empirical consequences of this and other differences between the two models of stochastic grammar remain to be investigated.

In terms of learning, both grammatical models can be learned with HG-GLA; when applied with Max-Ent-HG, this is stochastic gradient ascent (Jäger 2007). Given an input-output pair as learning data, the learner samples from the candidate set according to the probability distribution defined by the current weights; updating proceeds as outlined in 4.2. An attractive property of Max-Ent-HG is that there are existing proofs of convergence for associated learning algorithms that can be applied to the learning of variation (e.g. Fischer 2005); this has yet to be shown for Noisy HG. It is important to note that there is much room for further development of the learning algorithms for all of these models, since none of the convergence proofs apply when some of the structure of the learning data is hidden (see Boersma and Pater 2008 on Noisy HG, and Eisenstat 2009 Max-Ent-HG).

4.4 Applications to Dialectical Differences

To test the ability of these models of grammar to encode a range of probability distributions, we submitted distributions of *t/d*-deletion matching those from each of the dialects in (10) to learners implemented in Praat. The learners operated with stochastic OT (St-OT), Noisy HG (N-HG), and (ME-HG) grammars, using the OT-GLA and HG-GLA learning algorithms outlined above. For the Noisy HG learner, a non-negativity condition on weights was imposed in evaluation: if the ranking value (post-noise) was less than zero, it was replaced by zero (this is termed Linear-OT in Praat; see Keller 2000, 2006). We will discuss the motivation for the non-negativity condition shortly. The results are presented in (23). For each dialect, the top row indicates the observed proportion of deleted instances of *t/d* in each environment (pre-vocalic = CtV, pre-pausal = Ct, and pre-consonantal = CtC). The following rows show the final state constraint values for each model of grammar, and the encoded probability distributions (estimated using Praat's "get output distributions" method). In all cases but one, all the grammars encode probabilities closely matching the observed frequencies. In fact, the different models match the probabilities so closely that a statistical comparison of closeness of fit is unnecessary. The deletion patterns are reflected in the values of the constraints: when /CtV/ has the lowest rate of deletion, Max-P-V has a higher value than Max-Fin, and when /Ct/ has the lowest rate of deletion, the relationship is reversed.[7]

(23) Observed and learned *t/d*-deletion rates for different dialects of English

		*Ct	Max-P-V	Max-Fin	Max	CtV	Ct	CtC
AAVE (Washington, DC)						*0.29*	*0.73*	*0.76*
	St-OT	101.0	102.3	96.8	99.0	0.29	0.73	0.76
	N-HG	101.0	5.8	–1.5	97.2	0.29	0.73	0.77
	ME-HG	100.6	2.1	0.2	99.4	0.30	0.74	0.77
Chicano						*0.45*	*0.37*	*0.62*
	St-OT	100.4	99.7	100.6	99.6	0.45	0.37	0.62
	N-HG	100.4	1.0	1.8	99.6	0.43	0.36	0.60
	ME-HG	100.2	0.7	1.0	99.8	0.44	0.36	0.61
Jamaican						*0.63*	*0.71*	*0.85*
	St-OT	101.4	100.0	99.2	98.6	0.63	0.70	0.84
	N-HG	101.5	1.7	0.8	98.5	0.63	0.70	0.85
	ME-HG	100.9	1.2	0.8	99.1	0.64	0.73	0.85
Trinidad						*0.21*	*0.31*	*0.81*
	St-OT	101.2	103.4	102.5	98.8	0.21	0.31	0.80
	N-HG	101.2	5.2	4.1	98.8	0.21	0.31	0.80
	ME-HG	100.7	2.8	2.2	99.3	0.21	0.32	0.81
Tejano						*0.25*	*0.46*	*0.62*
	St-OT	100.4	101.9	99.6	99.6	0.25	0.46	0.62
	N-HG	100.3	1.5	0.7	99.7	0.25	0.47	0.62
	ME-HG	100.4	3.2	0.7	99.6	0.27	0.46	0.63
Tejano′						*0.62*	*0.46*	*0.25*
	St-OT	99.8	–6511.3	–523.2	100.2	0.45	0.45	0.45
	N-HG	99.8	–6382.1	–735.2	100.2	0.44	0.44	0.44
	ME-HG	99.4	–1.6	–0.8	100.6	0.61	0.42	0.24

The one case in which the learners did not all succeed in probability matching is labeled "Tejano-prime." This distribution was created by trading the proportions of deletion between pre-consonantal position and pre-vocalic position from real Tejano. The result is a pattern that exists in no known dialect: lowest frequency of deletion in pre-consonantal position. Stochastic OT and Noisy HG were unable to capture this pattern. For stochastic OT, as in the POC theory, this is because no ranking of the constraints yields deletion in only pre-consonantal position. Since stochastic OT produces a probability distribution over rankings, its restrictions on relative rates of variable processes have the same basic character as those of

the POC theory. Turning to the Noisy HG result, we note first that a categorical version of HG that restricts weights to positive values generates the same five languages as OT.[8] Since Noisy HG produces a probability distribution over weightings, in cases like these, where the OT and HG typologies converge, it also yields the same basic restrictions on relative rates of variable processes as POC and stochastic OT. This result is dependent on banning negative weights for the constraints, since a constraint that is negatively weighted will *prefer* the structure that violates the constraint. For the N-HG grammar for Tejano', the negatively weighted constraints have no effect on evaluation. The effect of negative weights is illustrated by the ME-HG result version that we used had no non-negativity restriction (this is not a necessary property of Max-Ent-HG models *per se*, since they can also be limited to positive weights). It was thus able to find a weighting that disprefers pre-consonantal deletion, by rewarding deletion in the pre-vocalic and phrase-final positions. Since constraint violations are marked with negative numbers, if the constraint weight is negative, then the product of the weight and the number of violations results in a positive increase of the Harmony of the candidate.

4.5 Variable Rules

As we pointed out earlier, Labov's (1969) variable rules notation specifies elements of the context of a rule as affecting the probability of application of the rule. We now show how this probability is calculated, and compare this probabilistic theory of grammar to the constraint-based ones just discussed. In a standard categorical rewrite rule of the form A → B / C __ D, A is changed to B every time that it occurs in the context C __ D, and only then. Whenever all the components in the structural description of the rule (C __ D) are present, the likelihood of rule application is 1.0, and whenever any of these components is absent, the likelihood of application is zero. Labov introduced the notion of weighting the components in the structural description such that each component contributes to the likelihood of rule application. Under this interpretation of rewrite rules, it becomes possible to say, for instance, that the presence of C __ in the context increases the likelihood of rule application by some specific factor, and similarly for the presence of __ D. The rule can now apply even if neither C __ nor __ D are present, and the likelihood of application can take on any value between zero and 1.0.

Several mathematical models have been proposed over the years for relating the observed application rate of some rewrite rule to the presence/absence of different components of the rule's context (Cedergren and Sankoff 1974; Rousseau and Sankoff 1978; Guy 1993; etc.). The one that has become the standard in the field, and that is implemented in the widely used software packages of VarbRul and Goldvarb, performs a multivariate stepwise logistic regression over observed token counts (Paolillo 2002: 177; Sankoff, Tagliamonte, and Smith 2005). In this analysis, application/non-application of the rule is taken as the dependent variable, and different factors hypothesized to influence the probability of application are taken as independent variables. Given a corpus of observed tokens to which

the rule could apply, and in which each token is coded for application/non-application of the rule, as well as for the value for each of the independent variables that hold of the specific token, VarbRul/Goldvarb estimates the contribution that each independent variable makes to the probability of rule application, using a maximum likelihood algorithm.[9] In this model, the probability that some rule will apply is given by the expression in (24). In this expression, p_0 is the "input probability", or the probability that the rule will apply independent of any of the contextual factor variables. The different independent variables are represented by 1 to n, and p_1 to p_n are then the contribution that each of these variables makes to the probability of rule application, as determined by the maximum likelihood algorithm. Stated in more concrete terms, an underlying form like /wɛst/ is subject to a variable deletion rule and therefore has two possible surface forms [wɛst] and [wɛs]. The expression in (24) defines a probability distribution over the two possible surface forms with the probability of the surface form to which the rule has applied ([wɛs]) being p, and the probability of the form to which the rule has not applied being $(1 - p)$. These probabilities depend on the values of p_0 to p_n.

(24) Probability of rule application in the variable rule framework

$$p = \frac{p_0 \times \ldots \times p_n}{[p_0 \times \ldots \times p_n] + [(1 - p_0) \times \ldots (1 - p_n)]}$$

To provide an illustration of this model that can be easily compared with those of the grammar models discussed above, we created a corpus for the Tejano data (Bayley 1995) that we also discussed in sections 3.2 and 4.4 above. Since we did not have access to Bayley's original corpus, we made the assumption that each of the three contexts (pre-consonantal, pre-vocalic, and phrase-final) appears 100 times. In our corpus, pre-consonantal position was marked with deletion (application of *t/d*-deletion) 62 times, phrase-final position 46 times, and pre-vocalic position 25 times, in accordance with the deletion rates that Bayley reports for these three contexts. We then submitted this corpus to Goldvarb X – the most recent version of the software package developed by David Sankoff to implement the statistical analysis described just above (Sankoff, Tagliamonte, and Smith 2005). The output generated by Goldvarb X is given in the first row of (25); variables with weights greater than .5 increase the probability of rule application, and lower weights decrease it. As explained above, p_0 is the input probability. Since we coded our data for only one independent variable (the following context), there are values only for p_1 in addition to the input probability. The expected deletion rates in the three contexts can now be calculated by substituting the value for p_0 and the appropriate value for p_1 into the equation in (24). For instance, to calculate the expected deletion rate in pre-vocalic position, we substitute .44 for p_0 and .30 for p_1. The resulting formula is shown in (26). Solving for p in this formula gives .25, which thus specifies the expected application rate of the *t/d*-deletion rule applying in pre-vocalic position. Substituting .52 for p_1 and solving for p gives

.46 as the expected deletion rate in pre-pausal context, and substituting .68 for p_1 gives .62 as the expected deletion rate in pre-consonantal context.

(25) Goldvarb X outputs[10]

	Factor weights				Observed and expected deletion rates					
					CtV		Ct		CtC	
	Input	CtV	Ct	CtC	O	E	O	E	O	E
Tejano	.44	.30	.52	.68	25	25.03	46	46.00	62	61.97
Tejano'	.44	.68	.52	.30	62	61.97	46	46.00	25	25.03

(26) Expected probability of *t/d*-deletion in pre-vocalic position in Tejano English

$$p = \frac{.44 \times .30}{[.44 \times .30] + [(1 - .44) \times (1 - .30)]} = .25$$

Variable rules, like stochastic versions of OT and HG, define a probability distribution over the possible surface forms for a given underlying representation. One difference between these two approaches is that OT and HG models are usually assumed to be restricted in the types of pattern that they can express by their use of a universal constraint set (though cf. Boersma 1998, Hayes, Zuraw, Síptár, and Londe 2009). We illustrated this aspect of these models in section 4.4 with the case of Tejano' – a dialect similar to actual Tejano except that the deletion rate in pre-vocalic and pre-consonantal position were inverted. Because the constraint set did not include a constraint blocking deletion in pre-consonantal position, the greater frequency of deletion in pre-vocalic position in Tejano' could not be modeled (with strictly positive weights in HG). Perhaps unsurprisingly, as shown in (25), our application of the variable rule theory was able to match the Tejano' distribution just as well as it could match Tejano. Paolillo (2002: Chapter 10) provides an explicit comparison of partially ordered constraint theorys (POC) with the variable rules model by reanalyzing Anttila's (1997a) Finnish data, and defends the lack of substantive restrictions on variable rules. The question of whether and how substantive restrictions on phonological processes should be encoded is of course a generally controversial issue in phonological theory; see Hale and Reiss (2000b) and Hayes, Kirchner, and Steriade (2004) for two poles of the debate (see further Hansson, this volume; see also Odden, this volume, on rules and constraints). It is worth noting, though, that the distinction between plausible and implausible phonological rules has often guided practice in formulating variable rules analyses; see especially Labov (2004) for discussion.

Another difference between research using variable rules and stochastic OT/HG is that social factors are usually included in variable rule analyses, but rarely, if ever, figure in analyses using the constraint-based frameworks. In variable rule

theory, a variable corresponding to a social factor like register can appear alongside linguistic variables like phonological context. For example, if *t/d*-deletion is more frequent in an informal register in Tejano English, this can be expressed with a second independent variable p_2. To illustrate, we show in (27) the result if this variable has the value .70 for the informal register, which raises the probability of rule application to .44. If the formal register had value 0 for this variable, it would have the .25 deletion rate calculated in (26).

(27) Expected probability of pre-vocalic *t/d*-deletion in a hypothetical informal register of Tejano English

$$p = \frac{.44 \times .30 \times .70}{[.44 \times .30 \times .70] + [(1 - .44) \times (1 - .30) \times (1 - .70)]} = .44$$

Some research in OT has proposed accounts of style/register differences. For example, van Oostendorp (1997) proposes that increasingly formal registers have increasingly high rankings of faithfulness constraints (see also Itô and Mester 2001a). In this proposal, however, each register is associated with its own categorical grammar, and variation results only from the selection of a different register and grammar. It is hard to see how such a model could capture the observed differences in frequency of deletion observed across languages/dialects, as in the English *t/d*-deletion case.

Boersma and Hayes (2001: Appendix C) suggest an approach to stylistic differences in stochastic OT in which the ranking value of a constraint is modified by a term expressing the degree to which that constraint's value changes in a given style. Concretely, they propose that *Style* is a real-numbered variable ranging from 0 to 1 (with 1 the formality maximum), and that *style-Sensitivity* is a constraint-specific variable that can take on positive and negative values. These variables are multiplied, and the result is added to the constraint's ranking value. In HG, this proposal would result in the Harmony equation in (28).[11]

(28) A candidate's Harmony in a style-sensitive HG

$$H = \sum_{k=1}^{K} (w_k + style \cdot styleSens_k) \cdot s_k$$

Boersma and Hayes do not provide a learning algorithm for the stochastic OT version of this model. In HG, at least, such algorithms do seem to be available. Given a version of HG that defines probability distributions over candidate sets, along with learning data annotated for style, one learning objective would be to find the values of the variables in (28) that minimize the difference between the observed distributions and the expected ones, that is, that minimize error. This sort of numerical optimization problem can be solved by a range of algorithms. The practical question of how to find weights for an analysis thus seems easily resolvable; a model that makes predictions about learning paths would be more challenging to construct.

We see the development of this and other variants of HG that include social variables as a particularly promising direction for further research. The development of these models could build on sociolinguistic research that examines the manner in which social variables impact the application of variable rules.[12] In the probabilistic model proposed by Cedergren and Sankoff (1974), as well as in Labov's (1969) earlier additive model and the later Goldvarb implementations, factors are independent. In terms of social and phonological factors, this predicts that the relative contribution of phonological factors cannot vary across registers. For example, the standard variable rules model cannot accommodate a variety of English that had the lowest rate of *t/d*-deletion pre-vocalically in one register, and pre-pausally in another. Although we do not know if such a dialect in fact exists, recent research suggests that the assumption of independence may be too strong (Bayley 2002: 130–134; Lin and Guy 2005). The model in (27) would allow for such non-independence, in that any subset of the constraints can be made sensitive to style (e.g. MAX-FINAL, but not MAX-PRE-V). A more restrictive model would make style sensitivity uniform across constraints, and/or limit the ways that constraint weights can be altered across registers, as in van Oostendorp's (1997) proposal; see, relatedly, Coetzee (2009a, 2009b). Further empirical research and development of these and other models is needed to choose between them.

5 Lexically Conditioned Variation

We now return to the main rhetorical thread of this chapter: the argument that variation cannot be left to "late phonology." In the last section, we discussed cases of phonological variation from English and other languages that display a characteristic of an "early" phonological process: sensitivity to morphological category. In this section, we discuss evidence that variable processes also show another purported diagnostic of early processes: sensitivity to lexical idiosyncrasy. This demonstration serves three purposes. First, it cements the case that it is insufficient to relegate variation to late phonology (or phonetic implementation), insofar as late phonology operates strictly on the output of early phonology, and is disallowed access to lexical representations. Second, it serves to raise some issues for the accounts of variation in OT and OT-like models discussed in the previous section. Third, it serves as a springboard for our discussion of directions for further development of constraint-based theories of phonological variation. Much of the data that we discuss here has formed the basis of recent arguments that a standard assumption about lexical representation in generative phonology is inadequate, that instead of each morpheme being phonologically represented in terms of a single abstract underlying form, each one is associated with a set of phonetically detailed exemplars (Bybee 2001; Pierrehumbert 2001a, 2002; see Ladd, this volume and Ernestus and Baayen, this volume). We start by examining the consequences of these data for theories of phonological grammar, in particular the constraint-based ones discussed in Section 4, before moving on to briefly discuss the implications for theories of the phonological lexicon.

5.1 Lexically Conditioned Variation in English

One of the cases of lexically conditioned variation discussed by Bybee (2001) has figured prominently in generative phonology ever since its conception: English secondary stress and vowel reduction. As Chomsky and Halle (1968) and many subsequent investigators have noted, words of the same phonological shape often have different secondary stress patterns. If we follow Chomsky and Halle and take Kenyon and Knott's (1953) pronunciation dictionary as our data source, words fall into three classes: a syllable of a particular type in a particular position can be consistently stressed (have a full vowel), consistently stressless (have a reduced vowel), or vary between stressed and stressless. In (29), we provide two examples from Pater (2000); see that paper for further discussion and references to earlier work.

(29) a. *Sonorant-final syllables that follow a heavy syllable and precede a stressed syllable*
Stressed: augmentation, condensation, importation, chimpanzee, incarnation, ostentation
Stressless: information, segmentation, transportation, Mozambique
Variable: advantageous, authenticity, condemnation

b. *Sonorant- and obstruent-final syllables in initial pretonic position*
Stressed: bandana, pontoon, bacteria, cognition, emporium, excursus
Stressless: Atlantic, admire, companion, confection, embrace, excursion
Variable: ambassador, Atlanta, Kentucky, September, sincere, obscene, accelerate

The data in (29) abstract away from important subregularities. For example, it seems that stressed category in the (a) cases is less well populated than in the (b) cases, and is usually dependent on the presence of a stress in the base form of a derived word. In addition, stresslessness in the (b) cases seems more productive in words with (historic) Latinate prefixes. Dealing with these subregularities would take us too far afield. For present purposes, we note just that the lexical idiosyncrasy cannot be explained away by the subregularities: derived words in (a) vary in whether they preserve the stress of their bases, and not all the cases of stress in this position are in derived words. For the cases in (b), it is not just words with Latinate prefixes that show reduction. Thus, any descriptively adequate account of these facts will have to accord a role for the lexicon in determining whether or not reduction takes place. Crucially, the lexicon does not fully determine whether reduction occurs: see, for example, the discussion of the categorical absence of stress on non-initial pretonic light syllables in Pater (2000).

It is of course possible that what Kenyon and Knott (1953) transcribed as variation was confined to inter-speaker variation, and did not include any genuine cases of within-speaker variation. However, we see it as highly likely that for at least some speakers, there are words like those in (29) for which there are two acceptable pronunciations, which are both produced in utterances that are in all

relevant respects identical. We also suspect that variation is underreported in Kenyon and Knott (1953).

English vowel reduction thus demonstrates that a variable process can display lexical idiosyncrasy in whether it applies or not. We now return to the case of English *t/d*-deletion to discuss evidence that the role of the lexicon in variation can be more fine-grained: that it can affect the frequency of application of a variable process. Thus, the determination of the probability of application of a phonological process must take into account not only the morphological and phonological factors discussed in the last section, but also the lexical item in question.

Like English vowel reduction and other lenition processes, *t/d*-deletion is more likely to apply to words with a higher usage frequency than words with a lower usage frequency (see especially Phillips 2006). Studies on *t/d*-deletion often leave out words like *just*, *went*, *and*, and *n't*, since these words typically show anomalously high deletion rates. As pointed out by both Bybee (2000: 70) and Patrick (1992: 172), these are words that are used with very high frequency. Motivated by this observation, Bybee reanalyzed the data collected by Santa Ana (1991) for Chicano English. She selected 2,049 tokens of words that end on /-Ct/ or /-Cd/ from Santa Ana's corpus. All these tokens were then divided into two groups based on their Francis-Kučera (1982) frequency. The "high frequency" group all appeared 35 or more times per million, and the "low frequency" group less that 35 times per million. As shown in (30), she found significantly higher deletion rates in the high frequency than in the low frequency words. See also Coetzee (2009a, 2009b) for similar evidence from different dialects of English.

(30) Rate of *t/d*-deletion in Chicano English

	Deletion	*Retention*	*% Deletion*
High frequency	898	752	54.4%
Low frequency	137	262	34.4%

Bybee (2001) also draws attention to the role that lexical frequency plays in the propensity for vowel reduction. As first noted by Fidelholtz (1975), frequent words are more likely to show reduction. Indications of this correlation can be glimpsed in the words in (29): *transportation* and *information* are more frequently used than *importation*, and *embrace* and *excursion* are more common than *emporium* and *excursus*. We do not take a position here on how or even whether the correlation between usage frequency and the frequency of the application of a variable process should be captured in a model of phonological grammar (see Coetzee 2009a, 2009b for two different approaches). At a minimum, however, these examples show that the probability of participation in a variable process is conditioned to some extent by lexical idiosyncrasy. Furthermore, even in a model that does take word frequency into account, there is almost certainly a residue of lexical influence that is independent of frequency (that is, of "exceptionality"). For example,

Coetzee (2009a) calculated the deletion *t/d*-deletion rate for several words in the *Buckeye Corpus* (Pitt *et al*. 2007).[13] Although he found a positive correlation between lexical frequency and deletion rate, there are many individual words that do not follow this general trend. The words *sound* and *friend*, for instance, have very similar frequencies in CELEX (Baayen *et al*. 1995) and are also phonologically and morphologically similar (both are monolsyllabic monomorphemes that end on /-nd/). However, they have very different deletion rates in pre-consonantal position, as shown in (31). The words *list* and *east*, which are also morphologically and phonologically similar (monosyllabic monomorphemes that end in /-st/), have very different CELEX frequencies, but more comparable deletion rates in pre-consonantal position. Lexical frequency, morphological status and phonological properties do not fully determine the probability of *t/d*-deletion, and accepting some lexical idiosyncrasy seems unavoidable.

(31) Deletion rate of selected words from the Buckeye Corpus

Word	*CELEX frequency*	*Log CELEX frequency*	*Deletion rate in Pre-Consonantal context*
sound	2,989	3.48	62
friend	3,087	3.49	81
list	1,360	3.13	56
east	2,409	3.38	59

We have focused here on the lexical conditioning of two English variable processes because they have been the subject of such careful scrutiny, and because it has been possible for us to directly observe their application relatively easily. However, we note that there are other well-studied cases of phonological variation that clearly show the role of lexical idiosyncrasy. Dutch has the exact same process of *t/d*-deletion, and this process in Dutch also has the same properties as in English – the process applies at different rates to lexical items that are identical in all relevant phonological properties and that have the same usage frequency (Goeman 1999; Goeman and van Reenen 1985; Hinskens 1992, 1996; Schouten 1982, 1984). For instance, the words *blaast* 'blow-3SG PRESENT' and *danst* 'dance-3SG PRESENT' have virtually identical CELEX frequencies (104 and 105, respectively), yet they differ quite substantially in their deletion rates: *blaast* undergoes deletion 16% of the time and *danst* 9% (Goeman 1999: 182; Phillips 2006: 65).[14]

Another process that provides a nice parallel to English *t/d*-deletion is French "schwa" deletion.[15] It similarly has some of the hallmarks of a late phonological process. Deletion of schwa from the initial syllable of a polysyllabic word (e.g. [sœmɛn] vs. [smɛn] *semaine* 'week') is variable and sensitive to rate of speech and style, is sensitive to phrasal context, and it seems to sometimes produce

between-category outcomes (Fougeron and Steriade 1997; Barnes and Kavitskaya 2003; cf. Côté and Morrison 2007). However, Dell (1973/1980: 206) notes that deletion of schwas from word-initial syllables does have exceptions, and claims that careful study of the phonological properties of the exceptional and non-exceptional words reveals no "simple regularity" that predicts whether a word will be exceptional. Thus, we have another process that combines aspects of early and late phonology (see Walker 1996 for discussion of the challenges that schwa deletion poses for Lexical Phonology in particular). Dell also notes that the exceptional words tend to be rarely used or literary. Racine and Grosjean (2002) conducted a production study with speakers of Swiss French, and found a continuum of frequency of deletion across different lexical items. While they did find a correlation between frequency of use of the lexical items in a corpus and frequency of deletion in their experiment, they note that the correlation is not perfect. Thus, as with English *t/d*-deletion, the available evidence indicates that words can be idiosyncratically resistant, to various degrees, to the variable process of French schwa deletion. And as with English *t/d*-deletion, the wealth of research on French schwa deletion makes this an ideal empirical base for the further development of theories of variation.

5.2 *Lexically Indexed Faithfulness Constraints and Variation*

As well as being problematic for theories that deny the access of variable processes to the lexicon, the data we have just discussed are problematic for the constraint-based models of variation overviewed in Sections 3 and 4, insofar as they also provide no way for the lexicon to affect the application of a variable process. Here we discuss one solution that draws on an existing proposed elaboration of OT: the indexation of faithfulness constraints to individual lexical items. Lexically indexed faithfulness constraints were introduced into OT in Pater (2000) as a means of dealing with influences of the lexicon on English secondary stress like those illustrated in (29).

As a simple illustration of how this would work for *t/d*-deletion, we provide the result of another learning simulation, this time giving the learner a hypothetical set of frequencies of deletion for both *feast* and *most* in the three "external" contexts. For both words, frequency of deletion was greatest in pre-consonantal position, intermediate in pre-vocalic position, and least in pre-pausal position. In each position, the frequency of deletion was greater for *most* than for *feast*. The learner had a Noisy HG grammar, and we elaborated our constraint set by creating versions of the faithfulness constraints specific to each of these lexical items. The results, provided in (32), show that the learner succeeded in probability matching for the distributions of deletion for both words. The weights of the faithfulness constraints specific to *feast* are higher than those specific to *most*, so that the grammar produces lower rates of deletion for *feast* than *most*.

(32) Results of a learning simulation with lexically specific constraints

*Ct	Max-Pre-V-/ feast/	Max-Fin-/ feast/	Max-/ feast/	Max-Pre-V-/ most/	Max-Fin-/ most/	Max-/ most/
100.71	1.18	3.58	100.07	1.16	3.27	99.22

	feast_V	feast	feast_C	most_V	most	most_C
Probability of deletion in learning data	*0.40*	*0.20*	*0.60*	*0.50*	*0.30*	*0.70*
Probability of deletion in learned grammar	0.40	0.20	0.59	0.50	0.30	0.70

Lexically specific constraints are a departure from standard generative assumptions about the nature of lexical representation, and its interaction with the grammar, in that they blur the distinction between the lexicon and the grammar, allowing them to interact in novel ways.[16] Such an interactive theory of the phonological grammar and the lexicon seems to be required to allow the lexicon to impact the application of variable processes. There is also increasing evidence from psycholinguistic research for richer interactions between the lexicon and perception and production than standard models permit; see Baayen (2007) and Ernestus and Baayen, this volume, for reviews.

There are a number of alternatives to lexically specific constraints that deserve to be explored, especially in conjunction with weighted constraint and other stochastic theories of grammar. A theory of the phonological lexicon with numerically enriched representations would likely resemble existing connectionist models; this resemblance could be exploited in the development of accounts of processing and learning (see e.g. Goldrick 2008). Another way to develop a highly interactive theory of lexical representation and phonological grammar would be to build on research on probabilistic approaches to language. If, as Pierrehumbert (2001a) has proposed, the lexicon defines a probability distribution over phonetically detailed forms of each morpheme, then this probability distribution could be made to interact with the probabilities for different contexts given by the phonological grammar.

Finally, in the context of a discussion of the nature of the representations manipulated by a phonological grammar, we must return to the issue with which we began this chapter. The view that variation belongs in the late phonology is consistent with the observation that variable processes often yield between-category outcomes. As we mentioned above, this observation holds for the case that we have paid the most attention to, English *t/d*-deletion (Browman and Goldstein 1990). The analyses that we have discussed do not generate these outcomes, being limited to categorical deletion or preservation of the word-final consonant. Our discussion emphasized the advantages of numerically weighted constraints for modeling the probabilities of deletion across dialects. It is likely that weighted constraints are also better suited than ranked constraints for generating phonetically

gradient outcomes (though cf. Boersma 1998), and may well help in this way to flesh out the accounts of *t/d*-deletion we have sketched (see Flemming 2001b and Kirchner and Moore 2008 on unified HG accounts of phonetics and phonology).

6 Conclusions

A theory of generative phonology that produces a probability distribution over outputs for an underlying representation has existed for nearly 40 years. Research using Labov's variable rules model has generated a wealth of information about the nature of phonological variation, of which we have discussed only a small sample. We have drawn on some of these data to show that variable processes are necessarily phonological, in that they are conditioned by morphological and lexical factors. We have also highlighted the overlap between the phonological factors conditioning the application of categorical and variable processes. However, phonologists outside the sociolinguistic tradition have been reluctant to embrace variable rules. Some (but clearly not all) of this reluctance may be attributed to the fact that the probabilistic component of the theory is largely independent of the rewrite rule formalism. The situation is quite different for the constraint-based theories of phonological variation that we have surveyed in this chapter. In these, the core formal mechanism of constraint prioritization (by ranking or weighting), determines cross-linguistic differences in probability of process application, just as constraint prioritization determines categorical differences between languages in the original version of OT. That constraint-based models of probabilistic phonology are firmly grounded in the core formalisms of the framework bodes well for the continued placement of variation as a central topic of research in generative phonology.

ACKNOWLEDGMENTS

Thanks to Matthew Goldrick, Bernard Laks, Jeroen van de Wejer, and the editors of this volume for comments on an earlier draft. The preparation of this paper was partially supported by grant BCS-0813829 from the National Science Foundation to the University of Massachusetts Amherst.

NOTES

1 Labov (1989) actually gives the following hierarchy from consonants that are most likely to induce deletion to those that are least likely: /s/ > stops > nasals > other fricatives > liquids. However, Guy and Boberg (1997) show convincingly that what is really relevant is the number of features that the preceding consonant shares with *t/d* – the more shared features, the higher the deletion rate. See also Coetzee (2004: ch. 5) for evidence in agreement with Guy and Boberg.

2 An alternative account of this asymmetry, consistent with Labov's (1997) observation that the allophonic details do not support the resyllabification account, is that homorganicity blocks release in [tl] but not [tr], and that unreleased consonants have higher rates of deletion (thanks to Lisa Selkirk for discussion).
3 The data reported in this table come from the following sources: AAVE (Fasold 1972), Chicano (Santa Ana 1991), Jamaican (Patrick 1992), Trinidad (Kang 1994), Tejano (Bayley 1995).
4 The theories are not in a subset relation: see Anttila (2007) for abstract examples of patterns of variation that POC can generate but that cannot be produced by the standard version of stochastic OT.
5 We do note that with more phonetically detailed representations and constraints, it might be possible to create a stochastic OT system that yields only gradient devoicing for geminates in isolation, and categorical devoicing for geminates in the context of another voiced obstruent (see especially Kawahara 2006 on the data).
6 Hayes and Wilson (2008) present another application of the Maximum Entropy framework to phonology. They use a log-linear model to define a probability distribution over the space of possible words, that is, as a model of phonotactics.
7 All of the input files used in the learning simulations reported here are included in "coetzee-pater-variation.zip," which is available from the authors, or from http://people.umass.edu/pater/coetzee-pater-variation.zip.
8 Readers interested in verifying this result can submit the file "typology.txt" from "coetzee-pater-variation.zip" to OT-Help (Staubs, Becker, Potts, Pratt, McCarthy, and Pater 2010).
9 For a detailed discussion of the mathematical model underpinning variable rule analyses, see Paolillo's instructive and accessible study (Paolillo 2002).
10 The files that were used as input to Goldvarb X are included in the aforementioned "coetzee-pater-variation.zip."
11 As Edward Flemming (p. c.) points out, this approach has a precedent in the 'carefulness' weight of Lindblom's (1990) H & H theory of speech production, which controls the extent to which target undershoot is minimized.
12 Thanks to John McCarthy for discussion of the issues raised in this paragraph, and especially for pointing us to Lim and Guy (2005).
13 This corpus consists of phonetically transcribed audio recordings of sociolinguistic interviews with 40 speakers of the Columbus, Ohio, dialect of English. For more detail on the corpus, see Pitt *et al.* (2007). For more on how the deletion rates were determined, see Coetzee (2009a).
14 Goeman calculates the deletion frequencies from the "Phonological and Morphological Properties of Dutch Dialects" database. See Goeman and Taeldeman (1996) for more on this database.
15 The scare quotes are to indicate that the vowel is not typically a phonetic schwa. We cannot here do justice to the complexities of this phenomenon, nor to the voluminous literature on the topic. Our brief discussion is based mostly on Dell (1973/1980); see Durand and Laks (2000) for discussion of Grammont's classic treatment, Eychenne (2006) for a recent overview of the generative literature and OT analysis, Tranel (2000) for an earlier OT approach, and Kimper (2008) on schwa deletion in a serial version of OT.
16 Lexically specific rules used in Chomsky and Halle (1968) and much subsequent work (see Zonneveld 1978), which give a similar power to the lexicon, became unfashionable with the rise of autosegmental phonology in the 1980s.

14 The Syntax-Phonology Interface

ELISABETH SELKIRK

1 Introduction

The topic of the syntax-phonology interface is broad, encompassing different submodules of grammar and interactions of these. This chapter addresses one fundamental aspect of the syntax-phonology interface in detail: the relation between syntactic constituency and the prosodic constituent domains for sentence-level phonological and phonetic phenomena. Two further core aspects, which rely on an understanding of the first, are not examined here – the phonological realization (spell-out) of the morphosyntactic feature bundles of morphemes and lexical items that form part of syntactic representation[1] and the linearization of syntactic representation which produces the surface word order of the sentence as actually pronounced.[2]

Early observations in the context of generative grammar of the apparent effects of syntactic constituency on phonology indicate already that the presence or absence of various types of phonological phenomena at different locations within a sentence correlates with differences in syntactic structure. Chomsky and Halle 1968 observed the tendency for local maxima of prosodic stress prominence to fall on the rightmost constituent within a phrase: [[A sènator [from Chicágo]] [wòn [the làst eléction]]]. McCawley 1968b recognized that in Tokyo Japanese "initial lowering" – a LH tone sequence – appeared at the left edge of groupings that correlate (in part) with syntactic constituency. Selkirk 1974 reported that in French the absence of word-final consonant deletion before a following vowel, referred to as *liaison*, also correlates with syntactic structure, as seen in the pronunciation of the adjective *petit* with final *-t* or without it: [[Le petit_âne] [le suivait]] the little donkey him-followed "The little donkey followed him" *vs.*

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

[[Le petit] [[aime] [le Guignol]], the little one loves the Guignol, "The little one loves the puppet theater." Subsequent research has expanded our understanding of the types of phonological phenomena that may be domain-sensitive in this very general sense: the full set includes a broad range of markedness-driven tonal phenomena of the sort that are independently attested in studies of word-level tonology and a broad range of markedness-driven segmental phenomena. There has also has been considerable phonetic research testifying to domain-sensitivity in the phonetic interpretation of the sentence.

In the last 30 years or so, debate has persisted around a central basic question: What is the nature of the linguistic representation in terms of which domain-sensitive phenomena of sentence phonology and phonetics are defined? Does the syntactic representation alone itself provide the structure in terms of which these domain-sensitive phenomena are defined (Cooper and Paccia-Cooper 1980; Kaisse 1985; Odden 1987, 1994b, 1996, 2000; Elordieta 2007d; Tokizaki 2008; Wagner 2005, 2010; Pak 2008, Samuels 2009, among others)? Or are there domains for phonology and phonetics that are defined in terms of a distinct prosodic structure that forms part of the properly phonological representation of the sentence (Selkirk 1978 *et seq.*; Nespor and Vogel 1982, 1986; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; Hayes 1989b; Inkelas 1990; Truckenbrodt 1995, 1999; Ladd 1996, 2008; Shattuck-Hufnagel and Turk 1996; Elordieta 1998, 2007c; Frota 2000; Seidl 2001; Dobashi 2003; Kahnemuyipour 2003; Gussenhoven 2004; Prieto 2005; Jun 2005; Revithiadou and Spyropoulos 2005, 2009; Ishihara 2007, among others)? Is a mixed picture of domain-sensitivity required, in which some phenomena are defined in terms of syntax directly and some in terms of a syntax-influenced phonological representation (Kaisse 1985; Seidl 2001; Pak 2008), and, if there is a mixed theory, which sorts of phenomena are sensitive to which sorts of constituency?[3] There is no easy answer to these questions. In this chapter, I will review and revise arguments for a prosodic structure representation of phonological domains, a representation that is independent of syntactic constituency but related to it by a module of syntactic-prosodic constituency correspondence constraints. The arguments to be presented in favor of a place in the theory of grammar for a prosodic constituent representation of phonological domains are based on recent rethinking of the nature of prosodic structure and the nature of syntactic structure-prosodic structure correspondence.

It does seems likely that the vast majority of domain-sensitive phenomena of sentence phonology as well as all of domain-sensitive phonetic phenomena are defined in terms of a properly phonological prosodic structure representation of domain. For example, depending on the language, the right or left edge of specific prosodic domains (whether prosodic word, phonological phrase, or intonational phrase) may identify the locus of local prosodic stress prominence, or the locus of tonal epenthesis, or the locus of consonant epenthesis or deletion, or the locus of segmental neutralization, and so on. These phenomena are arguably driven by a pressure for surface phonological representations to respect general constraints on phonological markedness, as construed, for example, within Optimality Theory (Prince and Smolensky 1993; McCarthy and Prince 1993b *et seq.*). At the same

time, it does also seem likely that certain phonological phenomena, like that of French *liaison* (in particular as it involves inflectional endings), are best analyzed as being directly sensitive to morpho-syntactic structure[4].

In what one might refer to as the "standard theory" of prosodic structure, prosodic constituent representation is defined as a well-formed labeled tree or bracketing, but one which has two fundamental properties that distinguish it from syntactic constituent structure representations – the prosodic hierarchy and strict layering (Selkirk 1978/1981b; 1981a, 1986: Nespor and Vogel 1986; Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; and others). The *prosodic hierarchy* is the name for an ordered set of prosodic category types. These prosodic category types constitute possible node labels for prosodic structures and in the standard view are stipulated by phonological theory.

(1) Prosodic category types of a commonly posited prosodic hierarchy

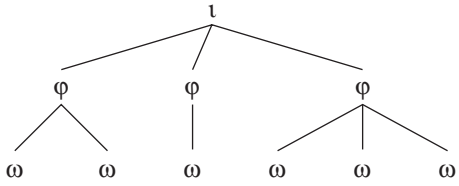
Intonational Phrase (ι)
Phonological Phrase (φ)
Prosodic Word (ω)
Foot
Syllable

In the standard theory no inherent relation is assumed to exist between the prosodic category types found in phonological representations and the category types of syntactic representation.

In this standard theory the nature of domination relations within a prosodic constituent structure is also stipulated by phonological theory. The *strict layer hypothesis* is the name given to the idea that a prosodic structure representation is strictly arranged according to the ordered set of categories in the prosodic hierarchy, as in (2). The strict layer hypothesis constitutes a purely phonological theory of the formal relations holding between constituents of the different prosodic category types in a prosodic structure.

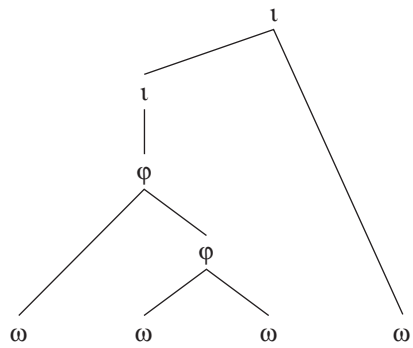
(2) The strict layer hypothesis

A constituent of category-level *n* in the prosodic hierarchy immediately dominates only a (sequence of) constituents at category-level *n*-1 in the hierarchy:



(Selkirk 1981a, 1986, 1995; Nespor and Vogel 1982, 1986; Pierrehumbert and Beckman 1988; Hayes 1989b; Inkelas 1990)

(3) A representation that violates the strict layer hypothesis:



In (3) are instances of configurations in which a constituent of a particular prosodic category type dominates another of the same category type (ι/ι and φ/φ); these are instances of *recursivity*. There is also an instance of a configuration in which a constituent of category-level *n* in the prosodic hierarchy immediately dominates a constituent of category-level *n*-2 (ι/ω); call this *level-skipping*.[5] These configurations both represent violations of the strict layer hypothesis. Strict layering predicts, among other things, that the edge of a higher-level prosodic category will always coincide with the edge of a category at the next level down in the prosodic hierarchy, with the consequence that the right edge of the sentence should always show the phonological properties of the right edge of phonological phrase and right edge of prosodic word, in addition to the properties of right edge of intonational phrase. If representations like (2) that do obey the strict layer hypothesis were the rule in phonology, as much earlier work contended, then phonological representations would indeed differ fundamentally from syntactic representations, which show configurations of the same general character as those found in (3). One first type of argument, then, for a non-syntactic, prosodic, representation of phonological domain structure was based on evidence that was taken to show that the domain structure for phonological and phonetic phenomena was indeed strictly layered, namely that it had formal properties distinct from that of syntactic constituent structure, as specified in (2) (see e.g. Selkirk 1978/1981b, 1981a, 1986; Nespor and Vogel 1986).

A second, related, argument for an independent prosodic constituency is based on the (putative) empirical generalization, found in this earlier literature, that phonological domain constituents may be systematically non-isomorphic to syntactic constituents. Early accounts attributed this putative non-isomorphism, or some of it, to the nature of the constraints relating syntactic structure and prosodic structure, which, in given syntactic configurations, were thought to result in mismatches between syntactic and prosodic constituency (see Nespor and Vogel 1986; Selkirk 1986, and discussion in Section 2.3).

A third type of argument for the distinctness of prosodic and syntactic structures is due originally to Nespor and Vogel 1986, who pointed out that non-syntactic

factors like speech rate may have an influence on phonological domain structure. Subsequent findings that phonological domain structure is also affected by phonological constraints on the weight or size of constituents or on their tonal properties (Selkirk and Tateishi 1988, Ghini 1993, Inkelas and Zec 1995, among others) provided further evidence for the non-syntactic character of phonological domains.

But evidence has been emerging that undermines the earlier claim made by the standard theory that representations of phonological domain structure systematically respect the strict layer hypothesis,[6] and the claim that constraints on syntax-phonology constituent correspondence produce a systematic non-isomorphism with syntactic and phonological constituents.[7] It turns out that the only argument for prosodic constituent structure that stands the test of time comes from non-syntactic influences on phonological domain structure. Section 3 of this chapter is devoted to elaborating this latter argument; it includes a review of phonological constraints on phonological domain structure, as well as a case study of prosodic phrasing in Lekeitio Basque (Section 3.2), which provides telling evidence for the role of properly phonological markedness-based factors in determining surface prosodic structure.

Section 2 addresses the nature of syntactic-prosodic constituency correspondence per se. Section 2.1 includes evidence from a case study of the Bantu language Xitsonga, which displays properties of prosodic constituent representation like those in (3) that go contrary to the strict layer hypothesis – namely recursivity and level-skipping. These properties are arguably the consequence of a new theory of the syntactic-prosodic constituency relation, one which calls for a match between syntactic and prosodic constituents (Selkirk 2006, 2009). An initial, pre-theoretic, version of a Match theory of syntactic-prosodic constituency correspondence is given in (4):

(4) Match theory of syntactic-prosodic constituency correspondence (to be refined[8])

i. Match clause
   A clause in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ι, in phonological representation.
ii. Match phrase
   A phrase in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it φ, in phonological representation.
iii. Match word
   A word in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ω, in phonological representation.

This set of universal Match constraints calls for the constituent structures of syntax and phonology to correspond; it predicts a strong tendency for phonological

domains to mirror syntactic constituents. The view to be argued for below is that the phonological constituent structure produced for individual sentences in individual languages, which includes a strong tendency to recursivity, is the result of syntactic constituency-respecting Match constraints like those in (4). Moreover, in identifying distinct prosodic constituent types (ι, φ, ω) to correspond to the designated syntactic constituent types, Match theory embodies the claim that, in the ideal case, the grammar allows the fundamental syntactic distinctions between clause, phrase, and word to be reflected in, and retrieved from, the phonological representation. At the same time, though, we will see that phonological markedness constraints on prosodic structure, if high-ranked, may lead to violations of Match constraints and produce instances of non-isomorphism between syntactic constituency and phonological domain structure. This independence of phonological domain structure from what is predicted by syntactic constituency provides the essential argument for the prosodic structure theory of this domain structure.

A general Match theory of syntactic-prosodic structure correspondence that encompasses ι-domains, φ-domains, and ω-domains has not before been proposed. But it has precursors in Ladd 1986 and Nespor and Vogel 1986 on intonational phrasing, and has re-emerged in more recent proposals within minimalist phase theory (Chomsky 2001) which hold that the Spell-Out domains of phases correspond to phonological phrases and/or that certain phrase types identifiable in terms of phase theory are spelled out as phonological phrases.[9] Further articulation of the Match theory of correspondence is found in Section 2.2, which also examines the implications of Match theory for the theory of prosodic structure itself. Match theory complements recent thinking by Itô and Mester (2007, 2010/11) on the nature of prosodic constituency representations in that it predicts that prosodic structure should display formal properties that bring it much closer to syntactic structure, including recursivity in constituency and a limited universal theory of prosodic category types. Section 2.3 is devoted to showing that Match theory is a better theory of syntactic-prosodic constituency correspondence than its predecessors, most notably the demarcative Alignment theory of Selkirk 1986, 1996 and the cohesional Wrap theory of Truckenbrodt 1995, 1999.

2 Syntactic Constituency and Phonological Domain Structure

2.1 A Case Study from Xitsonga

In Xitsonga, a Bantu language of northeast South Africa and Mozambique described and analyzed by Kisseberth 1994 and Cassimjee and Kisseberth 1998, the domain structure motivated by both tonal and segmental phenomena of sentence-level phonology shows a clear effect of syntactic constituency, but this domain structure also exhibits divergences from syntactic structure which are arguably the effect of prosodic structure markedness constraints. Kisseberth 1994 provides all the

empirical insights concerning the distribution of high tone spread and penultimate lengthening in Xitsonga sentences that are the basis for the analysis of phonological domain structure and domain-sensitivity offered here.[10] The data from Xitsonga is particularly valuable in that it provides evidence for the two distinct above-word levels of prosodic constituency that are grounded in the Match theory in (4), namely, the ι-domain and the φ-domain.

2.1.1 Penultimate Lengthening In Xitsonga, all and only vowels that are penultimate in the clause are long (V:). There is no lexical vowel length contrast; this penultimate lengthening is introduced by the phonology.[11] According to a purely syntactic approach, the clause would constitute the phonological domain for penultimate lengthening. But according to the prosodic structure hypothesis, and specifically Match theory, (4), the clause would correspond to an intonational phrase, or ι-domain, which itself would be the domain for penultimate lengthening:

(5)[12] a. [[ndzi-xav-el-a$_{Verb}$ [xi-phukuphuku] [fo:le]]]$_{clause}$ (K148)
1st.sg.Subj-buy-appl-FV class7-fool tobacco
'I am buying tobacco for a fool'

b. $_{ι}$(ndzi-xavela xi-phukuphuku fo:le)$_{ι}$

(Prosodic constituents internal to the ι will be shown when they are under discussion.)

Sentences like those below in (6) which contain a postposed subject show penultimate lengthening on the final word of the entire sentence as well as penultimate lengthening on the word preceding the postposed subject. According to Kisseberth (1994), the morphosyntax of such sentences argues that what precedes the subject is a clause itself.[13]

(6) a. $_{clause}$[$_{clause}$[[y-â:-j!á$_{Verb}$]]$_{clause}$ [n-gúlú:ve]]$_{clause}$ (K150–151)
Class9.subj-tense-eat-FV Class9-pig
'It's eating, the pig.'

b-i. $_{ι}$($_{ι}$(yâ:j!á)$_{ι}$ n-gúlú:ve)$_{ι}$ b-ii. * $_{ι}$(yâ:j!á)$_{ι}$ $_{ι}$(n-gúlú:ve)$_{ι}$

c. $_{clause}$[$_{clause}$[[vá-xáv-á$_{Verb}$ [tí-ho:m!ú]]]$_{clause}$ [vá:-nhu]]$_{clause}$
'They are buying cattle, the people are.'

d-i. $_{ι}$($_{ι}$(vá-xává tí-ho:m!ú)$_{ι}$ vá:-nhu)$_{ι}$ d-ii. *$_{ι}$(vá-xává tí-ho:m!ú)$_{ι}$ $_{ι}$(vá:-nhu)$_{ι}$

The nested prosodic ι-domain structure seen at the left in (6b-i, 6d-i) mirrors the embedded syntactic clause structure. That this recursive ι-domain structure would have to be posited in a prosodic account, rather than the alternative, sequential, ι-domain structure at the right in (6b-ii, 6d-ii), is shown by the ability of a lexical H tone that's final in the preceding clause to spread onto the postposed subject. As we will see below, such tone spreading is blocked at the left edge of both phonological phrases (φ) and intonational phrases (ι), so the postposed

subjects in (6) cannot be parsed as a φ or an ι. Assuming a prosodic structure account, the level-skipping recursive ι-domain structure here would be a first indication from Xitsonga that faithfulness to syntactic constituency leads to a violation of the strict layering that is assumed by the standard theory of prosodic structure.

Penultimate lengthening also appears at the right edge of syntactic phrases that are preposed in Xitsonga; it is found on any preverbal object, as in (7a, c, e). Any subject that immediately precedes the verb, like *hi-hontlovila* in (7c), does not show penultimate lengthening, but a subject preceding a preposed object does show it, as in (7e).

(7) a. $_{clause}$[[ti-ho:mú] $_{clause}$[[ndz-a-xa:v-a$_{Verb}$]]$_{clause}$]$_{clause}$ (K154–155)
'As for the cattle, I am buying.'

b-i. * $_{ι}$(ti-homú $_{ι}$(ndz-a-xa:va)$_{ι}$)$_{ι}$ b-ii. $_{ι}$(ti-ho:mú)$_{ι}$ $_{ι}$(ndz-axa:va)$_{ι}$

c. $_{clause}$[[ti-ho:mú] $_{clause}$[[hi-hontlovila] [x-!á-xá:v-a]]$_{clause}$]$_{clause}$
'As for the cattle, the giant is buying.'

d-i. *$_{ι}$(ti-homú $_{ι}$(hi-hontlovila x!áxá:va)$_{ι}$)$_{ι}$

d-ii. $_{ι}$(ti-ho:mú)$_{ι}$ $_{ι}$(hi-hontlovila x!á-xá:va)$_{ι}$) $_{ι}$

e. $_{clause}$[[n-sá:tí] $_{clause}$[[ti-n-gu:vu] $_{clause}$[[w-!á-xá:v-a]]$_{clause}$]$_{clause}$]$_{clause}$
'The wife, as for the clothes, she is buying.'

f-i. * $_{ι}$(n-sátí $_{ι}$(t-n-guvu $_{ι}$(w!á-xá:va)$_{ι}$)$_{ι}$) $_{ι}$

f-ii. $_{ι}$(n-sá:tí)$_{ι}$ $_{ι}$(t-n-gu:vu)$_{ι}$ $_{ι}$(w!á-xá:va)$_{ι}$

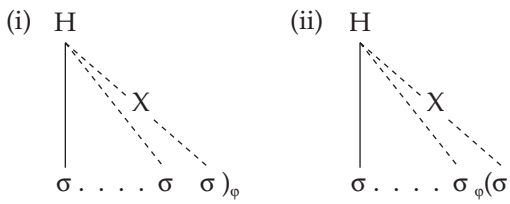
Kisseberth proposes that the preposed objects and subjects in (7) lie outside the clause, in the nested syntactic clause structure seen in (7a, c, e). However, the ι-domain structures that are motivated by the distribution of penultimate lengthening must be sequential, as in (7b-ii, 7d-ii, 7f-ii); they do not display the simple recursive embedding predicted by the syntax. If there were no right ι-domain edge at the edge of the preposed phrases, as in the ungrammatical prosodic structure parses in the recursive (7b-i, 7d-i, 7f-i), there would be no penultimate lengthening.

These Xitsonga data on penultimate lengthening in preposing structures show a certain divergence between syntactic structure and the phonological domain structure, given that the preposed phrases constitute ι-domains which do not correspond to syntactic clauses. That this divergence may have a source in some prosodic markedness constraint(s) is suggested by a comparison with the distribution of penultimate lengthening in the Bantu language Northern Sotho, a neighbor of Xitsonga in northeastern South Africa. Zerbian (2006, 2007) points out that penultimate lengthening appears sentence-finally in Northern Sotho, and also at the right end of the internal clauses in subject postposing cases, analogous to the Xitsonga cases in (5) and (6). But penultimate lengthening does not occur at the right edge of preposed phrases in Northern Sotho; the preposed phrases do not have the status of ι-domains, unlike in Xitsonga. Supposing that Northern Sotho and Xitsonga have the same clause-adjoining syntax for left-dislocations, as Zerbian argues, a non-syntactic explanation for the difference in the domain

structure for penultimate lengthening in the two languages would be required. The ι-domain structure of the Northern Sotho preposing construction is faithful to the syntactic clause constituency, unaffected by phonological constraints; it is predicted by the constraint Match Clause alone. But the Xitsonga preposing cases violate Match Clause in that they contain instances of prosodic ι-domains which do not correspond to syntactic clauses. Section 3.1.2 posits a markedness constraint Strong Start which would motivate the promotion of preposed phrases in Xitsonga to ι-domain status. In the grammar of Xitsonga, Strong Start would outrank Match Clause; the grammar of Northern Sotho, by contrast, would rank Strong Start lower than Match Clause.[14]

2.1.2 High Tone Spread Xitsonga is a tone language, and like many Bantu languages, has lexical H tone but no lexical L tone. High tone in Xitsonga may spread long-distance-fashion to the right onto toneless syllables. Limits on the extent of high tone spread in Xitsonga provide evidence for two clause-internal levels of phonological domain, one at the phrase level, referred to here as the φ-domain, and one at the word level, the ω-domain. Xitsonga high tone spread makes a very special contribution to the understanding of the relation between syntactic constituency and φ-domain structure in that the limits on the spread of high tone that are observed in Xitsonga allow both the left and the right edges of φ-domains to be diagnosed. In what follows we will see that the following generalization holds: a lexical high tone spreads rightward from its underlying position, but it is (i) blocked from spreading onto the final, rightmost, syllable of a φ-domain and (ii) blocked from spreading across the left edge of a φ-domain. This generalization is graphically depicted in (8).

(8) The limits on Xitsonga High Tone Spread diagnose right and left edges of φ



(9) shows the tonal patterns of verbs which constitute sentences on their own. The cases in (9a) consist of lexically toneless verb roots and lexically toneless prefixes, while those in (9b) contain the same toneless verb with a H tone prefix *vá*, third-person plural subject*:*

(9) a. [[ndz-a-tlomute:la$_{verb}$]]$_{clause}$ I-pres-fish [[ndz-a-ti:rha$_{verb}$]]$_{clause}$ I-pres-work (K139)

b. [[v-á-tlómúté:la$_{verb}$]]$_{clause}$ they-pres-fish [[v-á-tí:rha$_{verb}$]]$_{clause}$ (< vá-a-tirha) they-pres-work

c. $_{ι}(_{φ}(_{ω}($ v-á-tlómúté:la $)_{ω})_{φ})_{ι}$ $_{ι}(_{φ}(_{ω}($ v-á-tírha $)_{ω})_{φ})_{ι}$

In (9bc) the high tone of the subject prefix *vá-* spreads rightward through the verb but is blocked from spreading onto the final syllable of the sentence. Cassimjee and Kisseberth 1998 propose that a phonological constraint Non-Finality bans tone on a domain-final syllable: Non-Finality(D,H). The data from isolation forms in (9bc) does not indicate whether the domain D for Non-Finality is ι, φ, or ω, since the final syllable of this one-word sentence appears at the right domain edge at all these levels, as in (9c). But (10bc) shows that ω (prosodic word) is not the domain for Non-Finality: when the verb is not final in a phrase, as is the case in (10), the high tone spreads from the tone-bearing prefix throughout the verb and continues through a following toneless noun object up to the pre-final syllable of that noun.

(10)	a.	[[hi-tisa [xi-hontlovi:la]]]	[[ndzi-xava [nya:ma]]]	(K142)
		'We are bringing a giant.'	'I am buying meat.'	
	b.	[[vá-tísá [xí-hóntlóví:la]]]	[[vá-xává [nyá:ma]]]	
		'They are bringing a giant.'	'They are buying meat.'	
	c.	$_{\iota}(_{\varphi}(_{\omega}$(vá-tísá)$_{\omega}$ $_{\omega}$(xí-hóntlóví:la)$_{\omega}$ $)_{\varphi})_{\iota}$	$_{\iota}(_{\varphi}(_{\omega}$(vá-xává)$_{\omega}$ $_{\omega}$(nyá:ma)$_{\omega})_{\varphi})_{\iota}$	

The important fact is that H tone spreads onto and beyond the ω-final syllable of the verb.

The data from (11) (K148) show, more specifically, that a higher-than-word and lower-than-clause-sized domain, namely φ, is a/the domain for Non-Finality.

(11) a. [[ndzi-xavela [xi-phukuphuku] [fo:le]]]
'I am buying tobacco for a fool'
[[ndzi-xavela [mu-nhu] [ti-n-gu:vu]]]
'I am buying clothes for s.o.'

b. [[vá-xávélá [xí-phúkúphúku] [fo:le]]]
'They are buying tobacco for a fool.'
[[vá-xávélá [mú-nhu] [ti-n-gu:vu]]]
'They are buying clothes for s.o.'

c. $_{\iota}(_{\varphi}$(vá-xávélá xí-phúkúphúku)$_{\varphi}$ fo:le)$_{\iota}$
$_{\iota}(_{\varphi}$(vá-xávélá mú-nhu)$_{\varphi}$ ti-n-gu:vu)$_{\iota}$

High tone spread stops before the final syllable of the first object in (11b/c). This fact will be taken to indicate that the final syllable of the first object is at the right edge of a phrasal φ-domain and that Non-Finality holds of that φ-domain: Non-Finality(φ,H). The φ-final status of the last syllable of the indirect object *xi-phukuphuku* in these examples is indicated by the right φ bracket in (11c).

The constraint Non-Finality(φ,H), formulated as in (12a), rules out the appearance of a high tone on the final syllable of φ.[15] If we make the assumption that high tone spread is not itself domain-sensitive but has the completely general formulation in (12b), then it is the optimality theoretic ranking of H-Spread below Non-Finality(φ,H), as in (12c), that would serve to block the spread of high tone onto a φ-final syllable.

(12) a. Non-Finality(φ,H)

b. High Tone Spread (H-Spread)[16]

c. Non-Finality(φ,H) >> H-Spread

It is important to note in this connection that when a lexical tone originates on a word-final syllable in lexical representation and as a result of prosodic domain formation ends up in a φ-final position within the sentence, that φ-final H tone is indeed realized in surface representation, in violation of the Non-Finality constraint.[17] This is observed in the nouns *ti-homú* 'cattle' and *n-sátí* 'wife' in examples in (7) above; the first has lexical tone on its final syllable, the second has lexical tone on both final and pre-final syllables. In the examples in (7), where these nouns are final in preposed phrases, there is no rightward spread of that final lexical H tone. But in the example in (6), the final tone of the direct object *ti-homú* spreads onto the postposed subject. This example shows that high tone spread can take place from a word-final lexical tone across the right edge of a φ-domain and even across the right edge of a ι-domain, namely the instances of these that are found at the right of the embedded clause in (6). Given this, the preposing examples in (7) provide evidence that it is the left edge of the following ι-domain (or the left edge of any φ-domain that might appear at the beginning of that following clause) that is responsible for the blocking of high tone spread there (see Section 2.1.3). Additional examples support the hypothesis that the left edge of the φ-domain does block high tone spread.

Kisseberth observes that high tone spread may never penetrate the left edge of a noun phrase that contains a noun plus a modifier, while, as we saw above in (6), (10) and (11), it can spread into a noun phrase consisting of a single noun. This effect of noun phrase size or branchingness can be seen in the contrast in the distribution of high tone spread in the verb plus direct object constructions in (13i) and (13ii):

(13) (i) a. vá-súsá [n-gúlú:ve] [K157]
'They are removing a pig.'
b. $_{ι}(_{φ}$(vá-súsá n-gúlú:ve)$_{φ})_{ι}$
(ii) a. vá-súsá [n-guluve y!á vo:n!á]
'They are removing their pig.'
b. $_{ι}($ vá-súsá $_{φ}$(n-guluve t!á vo:n!á)$_{φ})_{ι}$

Note that the H tone, which originates in the verbal prefix *va*, extends only through the final syllable of the verb in (13ii),[18] and stops there. This H tone does not spread into even the first syllable of the two-word noun phrase.[19] In (13i), by contrast, the H tone extends two syllables into the object. If we assume the distinct prosodic φ-domain structures seen in (13i-b) and (13ii-b), we can attribute the

blocking of high tone spread in the latter case to the presence of the left φ-domain edge coinciding with the left edge of the multi-word direct object.

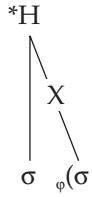
Further evidence of a difference in phonological domain structure for single-word and multi-word phrases in Xitsonga is found in cases of the blocking of rightward spread from a φ-final lexical high tone. Kisseberth reports that in configurations with a postposed subject or with a second object noun phrase of a ditransitive verb, if that noun phrase consists of a noun plus a modifier, a final lexical H tone cannot spread from a preceding word into the noun phrase (K159). The hypothesis here is that failure of high tone spread onto a multi-word second object or postposed subject diagnoses the presence of a left edge of φ, while presence of high tone spread into a following single-word noun phrase tells us that in such a case the noun phrase is not itself parsed as a φ-domain. The postposing examples above in (6) show spreading of a final lexical H tone from a clause-final word onto a following single-word postposed subject. The example in (14) shows high tone spread of the final lexical H tone in an indirect object medial noun phrase into the following single-word second object phrase *ti-n-gu:vu*.

(14) a. ndzi-nyíká [mu-nw!í] [tí-n-gú:vu] (K149)
b. 'I am giving the drinker clothes.'
c. $_{\iota}(\ _{\varphi}(\ \text{ndzi-nyíká mu-nw!í}\ )_{\varphi}\ _{\omega}(\text{tí-n-gú:vu})_{\omega}\ )_{\iota}$

In such cases, it is assumed that the single-noun phrase is not preceded by a left edge of φ, and the generalization holds that the left edge of a φ-domain blocks the spread of high tone, as stated in (8ii).[20]

In addressing the general question of domain-sensitivity in phonology, Itô and Mester 1999 propose a family of CrispEdge constraints which have the general property of blocking multiple linking of features across the edges of constituent domains. In this spirit, a constraint CrispEdgeLeft(φ,H) will be posited here; it rules out non-crisp-edge multiple linking of H tone across a left edge of φ-domain:[21]

(15) a. CrispEdgeLeft(φ,H)



b. CrispEdgeL(φ,H) >> H-Spread

The ranking of the domain-sensitive CrispEdgeLeft(φ,H) above H-Spread in the grammar of Xitsonga, as in (15b), will guarantee that high tone spreading into a φ-domain is not allowed.

It is necessary now to provide an analysis of the domain structure contained in the phonological representations that are evaluated by domain-sensitive

phonological constraints like Non-Finality and CrispEdgeLeft. A prosodic account of the φ-domain formation seen in cases like those just discussed would consist of two parts: (i) a syntactic-prosodic constituency correspondence constraint Match(Phrase,φ)[22] that calls for a syntactic phrase to correspond to a prosodic phrase φ in phonological representation and (ii) a prosodic markedness constraint – call it BinMin(φ,ω) – that calls for a φ to be minimally binary and thus consist of at least two prosodic words (Inkelas and Zec 1990b, 1995).[23] The size-dependent effects on domain structure would be the effect of BinMin(φ,ω). (Prosodic minimality constraints are also common at lower levels of the prosodic hierarchy, cf. Section 3.1.1). In Xitsonga a language-particular ranking of BinMin(φ,ω) above the phrase-matching correspondence constraint achieves the desired result, as shown in the optimality theoretic tableaux in (16):

(16) (i)

$_{clause}$[[verb [noun]$_{NP}$]$_{VP}$]$_{clause}$	BinMin(φ,ω)	Match(Phrase,φ)
a. $_{ι}$($_{φ}$(verb $_{φ}$(noun)$_{φ}$)$_{φ}$)$_{ι}$	*	
☞ b. $_{ι}$($_{φ}$(verb noun)$_{φ}$)$_{ι}$		*

(ii)

$_{clause}$[[verb [noun adj]$_{NP}$]$_{VP}$]$_{clause}$	BinMin(φ,ω)	Match(Phrase,φ)
☞ a. $_{ι}$($_{φ}$(verb $_{φ}$(noun adj)$_{φ}$)$_{φ}$)$_{ι}$		
b. $_{ι}$($_{φ}$($_{φ}$(verb noun)$_{φ}$ adj)$_{φ}$)$_{ι}$		*

In the optimal, grammatical, candidate (b) in (16-i.), a single-noun direct object is not parsed as a φ, in violation of Match Phrase; the Match-Phrase-respecting non-optimal candidate (16-i-a) violates the higher-ranked BinMin(φ,ω). But as the optimal candidate (a) in (16-ii) shows, a direct object can stand on its own as a φ if it contains more than one word, satisfying BinMin(φ,ω). In (16ii) both candidates satisfy BinMin(φ,ω); the optimality of (16-ii-a) is the effect of Match(Phrase,φ). The prosodic constituent structure of (16-ii-a) perfectly mirrors the syntactic structure, while that in (16-ii-b) does not. Note that (16-ii-a) shows a violation of strict layering, in that the verb stands external to the φ of its direct object, but does not itself have the status of a φ (as shown by the fact that tone may spread onto the final syllable of the verb in such cases).

It should be noted here that a Xitsonga sentence consisting of a verb plus two-word object and in addition a single-word postposed subject noun phrase would provide an instance of the prosodic structure in (3), which served to illustrate configurations which are in violation of the strict layer hypothesis, namely instances of recursivity and level-skipping. The general point here is that syntactic constituency has a central role in determining the phonological domain structure of a sentence, through the agency of Match correspondence constraints, but that the effect of syntactic constituency on that domain structure may, depending on the constraint ranking of the language, be mitigated by prosodic markedness constraints. This point will be elaborated in Section 3.

Consider next the prosodic parsing of a construction with verb followed by two single-word objects. The information available on Xitsonga syntax does not permit any decision on the details of the syntactic parsing of such ditransitive constructions, for which there are various possibilities, depending, for example, on whether the verb might have been raised to an inflectional head position, on whether the object(s) might have been raised to some higher specifier position(s), and so on. For this reason, it is impossible to be sure about the nature of the hypothetical Constraint X in tableau (17-i) which selects (b) as optimal from among the BinMin-respecting candidates (b) and (c); this remains a matter for future research.[24] What's important for the argument here is the significant difference in the φ-domain structure of the optimal (b) in (17i), with its two single-word objects, as compared to the optimal φ-domain structure (a) in (17-ii), with its two multi-word object noun phrases.[25]

(17) (i)

$_{clause}$[[verb [noun] [noun]]]$_{clause}$	BinMin (φ,ω)	Match (Phrase,φ)	Constraint X
a. $_{\iota}(_{\varphi}$(verb $_{\varphi}$(noun)$_{\varphi}$ $_{\varphi}$(noun))$_{\varphi})_{\iota}$	**		
☞ b. $_{\iota}(_{\varphi}$(verb noun)$_{\varphi}$ noun)$_{\varphi}$)$_{\iota}$		**	
c. $_{\iota}(_{\varphi}$(verb $_{\varphi}$(noun noun)$_{\varphi}$)$_{\varphi}$)$_{\iota}$		**	*

(ii)

$_{clause}$[[verb [noun adj] [noun adj]]]$_{clause}$	BinMin (φ,ω)	Match (Phrase,φ)
☞ a. $_{\iota}$($_{\varphi}$(verb $_{\varphi}$(noun adj)$_{\varphi}$ $_{\varphi}$(noun adj)$_{\varphi}$)$_{\varphi}$)$_{\iota}$		
b. $_{\iota}$($_{\varphi}$($_{\varphi}$(verb noun adj)$_{\varphi}$ $_{\varphi}$(noun adj)$_{\varphi}$)$_{\varphi}$)$_{\iota}$		*
c. $_{\iota}(_{\varphi}$(verb $_{\varphi}$(noun adj noun adj))$_{\varphi})_{\iota}$		**

The fact that two ditransitive structures like these that are identical in all but the word count of their object noun phrases should be spelled out with such different prosodic phrase structures is claimed here to be due to the subordination of Match(Phrase,φ) to BinMin(φ,ω) in the language-particular constraint ranking in Xitsonga.[26]

2.1.3 Typological Variation in Phonological Domain Structure and Domain-Sensitivity The optimality theoretic proposal being made here is that language-particular differences in prosodic structure that may be assigned to a sentence with a particular syntactic structure are a consequence of language-particular rankings of universal constraints on syntactic-prosodic constituency correspondence with respect to universal prosodic structure markedness constraints. Let us call this the interface theory of prosodic structure formation. But this interface theory of the interaction of constituency correspondence and prosodic markedness

constraints does not provide the whole story on cross-linguistic variation in domain structure and domain-sensitivity. Another source of cross-linguistic variation lies, of course, in language-particular differences in syntactic structure itself. For example, differences in the surface position of preposed constituents from one language to another – for example, whether they are located inside or outside the basic clause – can have consequences for the prosodic structure of the sentence.[27] And a third source of variation lies in the phonology proper, more specifically in the theory of domain-sensitivity, which was illustrated in the treatment of high tone spread in Xitsonga.

To further illustrate this latter point, we will complete the account of high tone spread in Xitsonga. It is necessary to tend to cases of the blocking of high tone spread which are not the consequence of the φ-domain structure of the language, but rather of its ι-domain structure. Consider the example of (6), where high tone spreads from the final syllable of the embedded clause through the postposed subject, but not onto the final syllable of the matrix clause. As argued above, the very possibility of tone spread onto the postposed single-noun phrase indicates that this phrase does not constitute a φ-domain. But if that postposed subject does not have the prosodic analysis of a φ, then the domain-sensitive constraint Non-Finality(φ,H) cannot be responsible for blocking the spread onto its final syllable. It must be an additional constraint Non-Finality(ι,H) that is responsible.

Given the general theory of the prosodic constituency hierarchy embodied in the Match correspondence constraints in (4), and given the idea that constraints like Non-Finality are defined with respect to prosodic constituents, it is of course to be expected that the universal repertoire of phonological markedness constraints should include a constraint Non-Finality(ι,H), in addition to a constraint Non-Finality(ω,H), and that their distinct effects should be attested in some language, as Cassimjee and Kisseberth (1998) propose. The fact that Non-Finality(ω,H) is not active in Xitsonga and that Non-Finality(φ,H) and Non-Finality(ι,H) are active can be ascribed to the optimality theoretic constraint ranking in (18).

(18) Non-Finality constraints and high tone spread in Xitsonga
Non-Finality(ι,H), Non-Finality(φ,H) >> H-Spread >> Non-Finality(ω)

It is to be expected that different rankings of H-Spread with respect to the various Non-Finality constraints would be found in other languages. For example, in a language where Non-Finality(ω) dominates H-Spread, no H tone could spread to a word-final syllable (and beyond). On the other hand, in a language in which H-Spread was dominated only by Non-Finality(ι,H), H-Spread would not be blocked at the right edge of ω or φ, but rather would have the capacity to spread across the span of the intonational phrase (if it weren't blocked by CrispEdgeLeft at the φ and ω levels.)

The prosodic structure theory of domain-sensitive constraints also makes available a family of CrispEdgeLeft constraints. The facts of Xitsonga are consistent with the ranking in (19), since high tone spread never penetrates the left edge of φ or of ι, but it passes through the left edge of ω:[28]

(19) CrispEdgeLeft and high tone spread in Xitsonga
CrispEdgeL(ι,H), CrispEdgeL(φ,H) >> H-Spread >> CrispEdgeL(ω,H)

By contrast, a language with a grammar where the constraint H-Spread was ranked below all three CrispEdgeLeft constraints would never permit spreading from one word to the following word. On the other hand, a language in whose grammar only CrispEdgeL(ι,H) dominates H-Spread would allow any left edge of φ or ω to be passed through by high tone spreading, up till the next left edge of ι (unless blocked by Non-Finality at the φ or ω levels). A theory of domain-sensitivity along the lines sketched here is a crucial component of a theory of cross-linguistic variation in the distribution of domain-sensitive phenomena within sentences.[29]

2.2 *The Nature of the Syntactic Constituency-Prosodic Constituency Relation*

In what follows, Sections 2.2.1 and 2.2.2 provide some needed elaborations of the Match theory of syntactic-prosodic constituency correspondence that was sketched in preceding sections. Section 2.2.3 reviews implications of Match theory for the theory of prosodic structure representation itself. In Section 2.3 Match theory is compared to other extant theories of the syntactic-prosodic constituency correspondence from both a theoretical and empirical point of view. Section 3 puts the focus on the prosodic markedness constraints that interact with Match constraints in prosodic structure formation.

2.2.1 Match Theory and Syntactic-prosodic Constituent Edge Correspondence

A highly restrictive general theory of the effects of syntactic constituency on phonological domains is presented in Selkirk (1986). Drawing on earlier proposals of Clements (1978) for Ewe and Chen (1987) for Xiamen which posited a relation between the edge of a syntactic constituent and the edge of a tone group in phonological representation, Selkirk (1986) proposes that only information about the *edges* of syntactic constituents of designated types is appealed to in constituency correspondence constraints. In its optimality theoretic instantiation in Selkirk (1996), this edge-based theory is referred to as the Alignment theory of the syntax-phonological phrasing interface, in the spirit of the generalized alignment theory of McCarthy and Prince (1993b). In this theory two distinct phrase-level constraints Align-R(XP,φ) and Align-L(XP,φ) are posited as part of the universal syntax-phonology interface constraint repertoire; the first calls for the right edge of a syntactic phrase XP to align with/correspond to the edge of a phonological phrase, the second calls for correspondence/alignment between left edges. The hypothesis was that languages could differ in which version of Align XP is responsible for prosodic phrasing patterns. Match theory shares with the edge-based theory of Selkirk (1986, 1996) this very limited appeal to the formal properties of syntactic phrase structure: Match(α,π) can be construed as a constraint requiring simply that *both* the right and left edges of a syntactic

constituent of a designated type α correspond, respectively, to the right and left edges of a prosodic constituent π. This dual-edge-matching entails a matching up of the constituents themselves. (These two theories receive a comparative evaluation in Section 2.3.2.)

Match theory, like Alignment theory, is a theory of universal constraints on the correspondence between syntactic and prosodic constituency in grammar. As the McCarthy and Prince (1995) theory of correspondence between distinct linguistic representations makes clear, correspondence is a two way street. In the case of correspondence between underlying (lexical) and surface phonological representation, an input-output correspondence constraint may require, for example, that a segment of the input has a corresponding segment in the output, thus disfavoring segmental deletion, while an analogous output-input correspondence constraint may require that a segment of the output correspond to a segment of the input, thus disfavoring segmental epenthesis. In the case of correspondence between syntactic and prosodic constituency, one set of correspondence constraints expresses the requirement that the edges of a syntactic constituent of type α must correspond to the edges of a phonological constituent of type π, while another set of correspondence constraints requires that the edges of a phonological constituent π correspond to the edges of a syntactic constituent α, as in (20). The first type – call them S-P faithfulness constraints – require that syntactic constituency be faithfully reflected in prosodic constituency, and the second – call them P-S faithfulness constraints – require that prosodic constituency be a faithful reflection of syntactic constituency. In the constraint schemata in (20), α is a variable over syntactic constituent types and π is a variable over their corresponding prosodic constituent types, as posited in (4):

(20) a. Match(α,π) [= S-P faithfulness]
The left and right edges of a constituent of type α in the input syntactic representation must correspond to the left and right edges of a constituent of type π in the output phonological representation.

b. Match(π,α) [= P-S faithfulness]
The left and right edges of a constituent of type π in the output phonological representation must correspond to the left and right edges of a constituent of type α in the input syntactic representation.

It does seem that both types of faithfulness are required in a theory of syntactic-prosodic constituency correspondence (see Selkirk 1996; Werle 2009).

Given the understanding of constraints on syntactic-prosodic constituency correspondence as faithfulness constraints, the proposal presented in the current chapter that these constraints interact, in language-particular fashion, with markedness constraints on prosodic representation is an entirely familiar one. In an optimality theoretic phonological component, a language-particular ranking of universal faithfulness and markedness constraints provides the basis for the phonological analysis of individual languages, while possible differences in language-particular rankings constitute a theory of cross-linguistic typology.

2.2.2 Constituency Correspondence Theory and Syntactic Category Types A central question for the Match theory of syntactic-prosodic structure faithfulness, as for any theory of the syntactic-prosodic constituency relation, concerns the choice of the syntactic constituents that figure in the correspondence constraints. What are possible α in the correspondence constraints schematized in (20)? The set of Match constraints proposed in (4) pare syntactic constituent types to the minimum, exploiting the notions clause, phrase, and word, which presumably play a role in any theory of morphosyntax. Assuming the general correctness of the claim embodied in (4) that a fundamental distinction between syntactic constituents of type clause, phrase, and word is made in constituency correspondence constraints, it remains to be determined just how to characterize in terms of current syntactic theory the notions "clause," "phrase" and "word" that are relevant to prosodic domain formation.

Consider first the clause-sensitive correspondence constraints Match(Clause,ι) and Match(ι,Clause). There are at least two notions of "clause" that come into play in syntax-prosodic structure correspondence, call them the "standard clause" and the "illocutionary clause." The standard clause is the constituent that is the complement of the functional head $Comp^0$. In modern syntactic theory, $Comp^0$, or simply C, is commonly assumed to introduce the canonical sentence, which consists of an explicit or implied subject, a predicate, and a locus for Tense:[30] $_{CP}$[$Comp^0$ [standard clause]]$_{CP}$. CP, and hence the standard clause, may be syntactically embedded, whether as a complement to a verbal or nominal head, or as a restrictive relative clause within determiner phrase, or in other positions. What is being called here the illocutionary clause is the highest syntactic projection of the sentence and carries its illocutionary force, which determines its appropriateness in a discourse context. Emonds (1976) termed this the root clause; Rizzi (1997) refers to this as the Force Phrase; it can be seen as an instance of the Potts (2005) comma phrase. The syntactic structure for this clause type will be assumed to be: $_{ForceP}$[$Force^0$ [illocutionary clause]]$_{ForceP}$. Parentheticals, non-restrictive relative clauses and other expressions of the type that Potts calls "supplements" may be embedded within the larger sentence, but they have the property that their meaning does not contribute to the "at issue" meaning of the surrounding sentence (Potts 2005). It seems reasonable to understand these as instances of embedded ForceP (see Kan 2009; Selkirk 2009). If $Force^0$ and $Complementizer^0$ do indeed form a natural class of functional heads, as is implied by the Rizzi (1997) proposal, then the general notion "clause" can be taken to designate a constituent that is complement to functional heads of the general complementizer class that includes these.

What are being called here illocutionary clauses are commonly observed to correspond to intonational phrases in phonological representation (see, e.g. Downing 1970; Nespor and Vogel 1986; Ladd 1986; Selkirk 2005; Dehé 2009 on English). It is less commonly observed, though apparently necessary, for standard clauses to correspond to intonational phrases. The Xitsonga cases of clauses embedded in dislocation structures would be instances of this sort, as would the cases where embedded standard clauses are reported to serve as a domain for phonological

phenomena in German (Truckenbrodt 2005), Huave and Luganda (Pak 2008), and Japanese (Selkirk 2009). Yet embedded illocutionary clauses appear to have a stronger tendency than embedded standard clauses to be prosodically parsed as ι-domains. In English, for example, the syntactic/semantic distinction between restrictive and non-restrictive relative clauses, or between parentheticals and basic clausal complements, is reflected in a difference in ι-domain structure (Ladd 1986; Selkirk 2005; Dehé 2009), with the former set off as ι-domains and the latter not. It is proposed in Selkirk (2009) that the theory of grammar distinguishes two instances of the Match Clause constraint, the more specific Match(illocutionary clause,ι) and the more general Match(clause,ι). Cross-linguistic evidence suggests that the former outranks the latter, perhaps universally.

Turning next to Match Phrase, the simple appeal to "phrase" in this constraint embodies the assumption made in previous theories of the syntactic-prosodic structure relation that distinctions between lexical projections NP vs. VP vs. AP are not relevant to this correspondence (Nespor and Vogel 1986; Selkirk 1986, 1996; Truckenbrodt 1999 and others). Selkirk (1986, 1996, 2005) and Truckenbrodt (1999) propose that the notion "maximal projection" (XP) from X-bar theory (Jackendoff 1977) is crucial to defining the notion "phonological phrase." This cross-categorial appeal to maximal projection predicts the sort of φ-domain organization that is arguably typical in languages: SVO sentences like NP $_{VP}$[Verb NP]$_{VP}$ are parsed into φ-domains as (NP)$_{φ}$ (Verb (NP)$_{φ}$)$_{φ}$; double object structures like $_{VP}$[Verb NP NP]$_{VP}$ are parsed as (Verb (NP)$_{φ}$ (NP)$_{φ}$)$_{φ}$.

But the notion XP needs to be further refined, since it is likely that lexical and functional phrasal projections – LexP and FncP – have to be distinguished (see discussion in Selkirk and Shen 1990; Truckenbrodt 1999). The functional vs. lexical distinction is important for syntactic-prosodic correspondence at the word level (Fnc^0 vs. Lex^0): lexical category words are standardly parsed as prosodic words (ω), while functional category words like determiners, complementizers, prepositions, auxiliary verbs, and so on – in particular the monosyllabic versions of these – are not (see e.g. Selkirk 1996; Werle 2009). Of relevance to the phrasing issue is the fact that a functional head may in some languages become prosodically enclitic to a preceding constituent that is not contained in the FncP that the Fnc word heads; this may be explained by assuming that a FncP is not delimited by φ-boundaries. Examples are the inclusion of a preposition from a following prepositional phrase into a prosodic word that includes the preceding verb in Shanghai Chinese (Selkirk and Shen 1990) and the inclusion of the determiner from a following determiner phrase into the prosodic word of the preceding verb in Kwakwala (Boas 1947; Anderson 1984, 2005) and Chamicuro (Parker 1999), or of a syllabic noun-class prefix from a following noun phrase, as in Xitsonga (Kisseberth 1994). The English forms *wanna* (< *want to* or *want a*), *gotta* (< *got to* or *got a*), *kinda* (< *kind of*), *shoulda* (< *should have*) are likely historical, or even synchronic, instances of this sort of thing. If instead of a general Match XP this correspondence constraint were limited to lexical categories,[31] then, on the basis of the syntactic structure $_{VP}$[Verb $_{FncP}$[Fnc NP]$_{FncP}$]$_{VP}$, the φ-domain structure $_{φ}$(Verb Fnc $_{φ}$(NP)$_{φ}$)$_{φ}$ would be predicted, namely a structure that would pose no obstacle to the prosodic encliticization of the Fnc

from the verbal complement onto the verb. Whether or not a language would as a result exhibit prosodic enclitization of the Fnc onto the verb – as in the prosodic structure $_{\omega}(_{\omega}($ Verb $)_{\omega}$ Fnc$)_{\omega}$ – would be driven by the ranking of relevant prosodic markedness constraints (see e.g. Werle 2009).

Proposals concerning the syntactic constituents relevant to prosodic constituency have been made within minimalist phase theory (Chomsky 2001). In that theory phases constitute stages in the derivation of syntactic structure and its interpretation; Spell-Out (phonological and semantic interpretation) completes each phase. The complement of any phasal head constitutes the domain of Spell-Out. It has been proposed that a Spell-Out domain itself corresponds to a prosodic constituent (see Adger 2006; Dobashi 2003; Ishihara 2007, for example). Two constituent types typically singled out for phasehood are νP and CP. The complement of phase head ν is the VP, or some functional projection containing the VP; it contains all the internal arguments of the verb.[32] The complement of the phase head C (complementizer) is an inflectional projection, typically the Tense Phrase; it contains all the material of the standard clause, namely subject, predicate, and tense-marking. It was proposed above that "complement of complementizer" does indeed identify the notion "clause" for the Match Clause constraint(s) under consideration here. But "complement of ν" does not on its own denote the full set of syntactic phrase types that are relevant to the correspondence with prosodic phrases, since these also include instances of lexical maximal projections like NP and AP. Moreover, Kahnemuyipour (2004, 2009) and Kratzer and Selkirk (2007) point out that in German all-new sentences, for example, not all lexical maximal projections have the stress and accent properties associated with φ-domains and they propose that those that do have the properties of φ-domains occupy the position of the highest phrase in the Spell-Out domain of a phase. The question of just how to syntactically define the set of syntactic phrases that may figure in the Match Phrase correspondence constraint(s) remains a question for further research. In what follows, for expository purposes, it will simply be assumed that Match Phrase stands for Match XP.

2.2.3 Implications of Match Correspondence Theory for Prosodic Structure Theory A new generalization about the nature of prosodic structure above the foot emerges from the review of Xitsonga sentence tonology in Section 2.1, namely that the prosodic constituent structure of a sentence is grounded in large part in the syntactic constituency of the sentence. It displays properties that are predicted by the Match theory of syntactic-prosodic constituency correspondence: (1) presence of systematic recursivity and level-skipping, in violation of strict layering, and (2) the presence of distinct prosodic domain types corresponding to clause, phrase, and word.

Note that the systematic recursivity of φ or ι domains that is predicted by the Match theory of the syntax-prosodic constituency correspondence provides the basis for accounts of known patterns in the phonetic realization of syntactically recursive structures such as have been studied by Lehiste (1973), Ladd (1986,

1988), Truckenbrodt and Féry (2005), and Wagner (2005, 2010), Féry and Schobö (2008), among others. Wagner (2005), for example, has shown that variant syntactic parsings of conjoined noun phrases such as (21ac), and their associated differences in semantic interpretation, are correlated with different patterns of final lengthening, with the greater lengthening occurring at the end of a more deeply embedded phrase:

(21) a. [Lysander and [Demetrius and Hermia]]
b. $_{\varphi}$(Lysander and $_{\varphi}$(Demetrius and Hermia)$_{\varphi}$)$_{\varphi}$
c. [Lysander and Demetrius] and Hermia]
d. $_{\varphi}$($_{\varphi}$(Lysander and Demetrius)$_{\varphi}$ and Hermia)$_{\varphi}$

The correspondence constraint Match Phrase converts syntactic representations like (21ac) into recursive φ-domain representations, which would minimally show the bracketing in (21bd). The phonetics would recognize this depth of φ-embedding in phonological representation and would assign the different values for final lengthening. Thus Match theory provides the basis for a prosodic structure-based account of the effects of syntactic structure recursivity on the phonological and phonetic interpretation of the sentence.

At this point, it is an open question whether prosodic markedness constraints that would enforce a flattening of phonological domain structure into a strictly layered representation are at play in grammar. Selkirk (1996) posits the existence of violable Non-recursivity and Exhaustivity constraints whose purpose is, respectively, to exclude instances of recursive and level-skipping prosodic structure. Markedness constraints like these which enforce strict layering (call them SLH constraints) will not be reviewed in Section 3. It remains a question for future research whether they are in fact needed in the theory of grammar. (Section 2.3.3 treats apparent counterexamples to a strong theory of prosodic structure formation which includes Match theory and excludes SLH markedness constraints.)

As for the repertoire of prosodic category types that figure in prosodic representations of sentences, the need for restricting this repertoire has been underlined in Itô and Mester (2007, 2010/11). In the proposal here, the small repertoire of distinct types ι, φ, and ω posited by Itô and Mester (and many others) derives from the theory of syntactic-prosodic constituency correspondence constraints. In the Match theory posited in (4), the fundamental types of syntactic constituent – clause, phrase, and word – are each identified with a distinct corresponding type of prosodic constituent in phonological representation: (clause,ι), (phrase,φ), (word,ω). What names are given to these distinct prosodic constituent types is immaterial. What is crucial is that there's a distinct prosodic constituent type that clauses are required to correspond to, referred to here as ι, or intonational phrase (following standard usage); there is the distinct prosodic constituent type that syntactic phrases are called on to correspond to, referred to here as φ, or phonological phrase; and there is the distinct type that words are required to correspond to, namely ω, or prosodic word.

Stated explicitly, this *syntactically grounded prosodic hierarchy hypothesis* holds that all and only the supra-foot prosodic category types that figure in syntactic-prosodic constituency correspondence constraints are defined as primitive prosodic category types in linguistic theory (see Selkirk 2005). There are two essential predictions of this hypothesis. One is that the syntactically grounded prosodic category types ι, φ, and ω are universal. This prediction follows from the hypothesized universality of Match correspondence constraints in the grammar. Unless some higher ranked prosodic markedness constraint(s) were to prevent the realization of constituents of one or more of these types in some language, they all should appear in every sentence of every language. The second prediction is that, where further types of prosodic category above the foot appear to be warranted in the phonology or the phonetics, they are in fact sub-types of the primitive, syntactically grounded category types ι, φ, and ω, as Itô and Mester (2007, 2010/11) propose. They point out that, given recursivity in φ-structure, for example, distinctions can be made between *maximal* φ – a φ not dominated by any other φ, *minimal* φ – a φ not dominating any other φ, and simple φ – the general case. They argue that phonological constraints may make appeal to these various sub-types.

The notion that prosodic category types above the foot are universally instantiated in the phonological representations of any language is not shared by all phonologists and phoneticians working within a prosodic structure framework. In a summary chapter in the volume *Prosodic Typology*, Jun (2005) writes "Languages seem to differ in how an utterance is rhythmically and prosodically organized. Based on the [autosegmental-metrical] model of various languages, some languages have only one prosodic unit above the word (e.g. Serbo-Croatian), while others have three (e.g. Bininj Gun-wok, Farsi) (443)." In the same volume Venditti (2005) posits two levels of prosodic organization for Japanese above the word, which she refers to as the accentual phrase and the intonational phrase. The former, smaller, phrasal unit has also been referred to as the minor phrase in work on Japanese; the latter is what has been referred to as the major phrase or intermediate phrase in other work on Japanese. What is explicitly claimed not to exist in the prosodic structure of Japanese sentences, in this and earlier presentations of Japanese sentence prosody in Venditti (1997) and Pierrehumbert and Beckman (1988), is a larger unit of prosodic structure of the sort typically referred to as intonational phrase. But Kawahara and Shinya's (2008) investigation of the prosody of Japanese sentences based on standard coordinate sentences and coordinate sentences with gapping has since shown the necessity of positing a prosodic category above the major/intermediate phrase whose edges coincide with clause boundaries, namely a prosodic category of the clause-grounded ι-domain level. They found that the ι-domain/clause edge exhibits final lowering, creaky voice, and pause not seen at a mere phonological phrase edge. (See Selkirk (2009) for a fuller discussion of clause and intonational phrase in Japanese.) Clearly, the hypothesis that prosodic category types are syntactically grounded and universal suggests an interesting program for further cross-linguistic research.

Paying systematic heed to syntactic structure is arguably a necessary component of cross-linguistic investigation of potential universals in the phonological and/

or phonetic interpretation of different levels of prosodic structure, given that it is not possible to establish what the prosodic levels of organization in a language are based only on the nature of the phonological or phonetic phenomena that are typically reported. Languages may differ in whether it is a phonological phrase or an intonational phrase that is the locus of a particular phonological or phonetic phenomenon (cf. 2.1.3). The important question whether there is in fact any cross-linguistic commonality on some phonological or phonetic dimension in the properties of intonational phrases or phonological phrases can be asked only if there is a theory that identifies independently of phonological or phonetic criteria which domain is an intonational phrase and which is a phonological phrase. An explicit theory of syntactic structure-prosodic structure correspondence like Match theory is just such a theory.

2.3 *Other Treatments of Syntactic-Prosodic Constituency Correspondence*

The year 1986 saw the publication of four influential works on the prosodic structure of sentences. Nespor and Vogel (1986) put forth a "relation-based" theory for defining phonological phrases as well as a Match theory of intonational phrases. Ladd (1986) presupposed a Match-based account of intonational phrasing. Selkirk (1986) argued for a single-edge-based theory of phonological phrasing. Beckman and Pierrehumbert (1986) assumed no particular relation between syntax and prosodic structure; their intention was to argue for general commonalities in prosodic structure organization and domain-sensitive phonetic interpretation in English and Japanese. Common to all but Ladd (1986) was the assumption that the prosodic structure of sentences conforms to the strict layer hypothesis (Selkirk 1981a. Ladd's contention in the 1986 paper and in more recent work has been, by contrast, that intonational phrases may be nested, in what has been termed here a recursive ι-domain structure. We saw above that a Match theory of the interface leads to potentially recursive ι-domain and φ-domain structure, with good results. The evaluation below of alternative theories of the syntax-phonological domain structure correspondence and of the data that motivated them will show that the typological predictions of Match theory are confirmed.

2.3.1 Nespor and Vogel (1986) on Phonological Phrasing Nespor and Vogel (1986) report that in Tuscan Italian gemination of a word-initial consonant following a stress-final word (*raddoppiamento sintattico (RS)*) is optionally possible between the head of a syntactic phrase and the first word of a following complement to that head if that complement phrase is non-branching: *Venderà questo leopardo* [will-sell [this leopard]] but *Prenderà qqualcosa* [will-take [something]]. Assuming that RS in confined to a φ-domain, they propose an interface prosodic phrase formation rule that that would restructure the Verb plus non-branching direct object in the second case into a single φ, but in the first case would leave the head and the following phrase separated into two φ: "A non-branching φ which is the first complement of X on its recursive side is joined into the φ that contains X

(173)." The actual domain-sensitive account of the phonology of RS that would be proposed with the current theory wouldn't necessarily be different, but the phrasing would be arrived at differently, with a simple Match Phrase constraint interacting with the prosodic constraint BinMin(φ,ω). (The Italian case is quite parallel to the one in Xitsonga that was illustrated above.) But Match Phrase is a double-edge-matching constraint; its language is minimalist; it avoids reference to a richer set of notions like linear order, adjacency, and relational notions like head of or first complement of that appear in statements in the relation-based theory. For this reason, all else being equal, the Match theory is to be preferred. The current theory of prosodic structure formation retains, though, the role for phrase size or branchingness in determining prosodic structure that was first recognized by Nespor and Vogel (1986); since then Ghini (1993), Inkelas and Zec (1995) and others have made the case that such branching effects in Italian should be understood as prosodic in character (see also Section 3.1.1).

2.3.2 Selkirk 1986, 1995: Align R/L(XP,φ) Like Match theory the single-edge-based theory of Selkirk (1986) hypothesizes a restricted appeal to syntactic structure constituency, except that it calls for the R or L edges of designated syntactic constituents to match up with edges of prosodic constituents, rather than for the entire constituent (the "node") to match up (via matching of both edges), as in Match theory. Selkirk (1986) makes a general proposal concerning syntactic-prosodic constituent correspondence according to which (i) depending on the language, either the R or the L edge of a designated constituent type in the syntax, for example, $]_{Xmax}$, must coincide with the edge of a corresponding prosodic constituent in phonological representation and (ii) the resulting constituency is governed by strict layering: "With the setting $]_{Xmax}$, an X^{max}-derived domain simply extends from one instance of $]_{Xmax}$ to another (or to the end of the sentence, if there is no further $]_{Xmax}$ [and similarly for the opposite setting-eos] (392)." From the outset, the edge-based theory presupposed the strict layer hypothesis, and, as a result, satisfaction of the edge correspondence constraint led to the formation of prosodic constituents that were non-isomorphic with the syntax. After the articulation of a generalized theory of alignment in McCarthy and Prince (1993a), the single-edge-based theory was dubbed the Align XP theory of phrasing.

It is important to observe that the single-edge-based Align XP theory is in fact underdetermined by the data from the languages which were originally taken to motivate it, since a Match XP analysis is equally consistent with the data. This is true of the cases of ChiMwiini, Xiamen Chinese and Tokyo Japanese, for example. Data from the extensive investigation of the distribution of vowel length in sentences of the Bantu language ChiMwiini reported in Kisseberth and Abasheikh (1974) is consistent with the Selkirk (1986) proposal that a right-end-of- X^{max} setting for the interface phrasing parameter derives the phonological phrasing manifested in the distribution of ChiMwiini vowel length. Kisseberth (2005) reports on further investigations of the distribution of vowel length as well as of the distribution of a right-edge phrasal tone in ChiMwiini (an otherwise non-tonal language), showing that the distribution of both these phonological phenomena converge on the

φ-domain structure that is predicted by the Align-R(XP,φ) analysis. The forward slashes provided in the (a) lines in (22) and (23) by Kisseberth (2005) informally represent the medial φ boundaries predicted by the single-XP-edge-based theory, while the (b) lines give the syntactic structure that Kisseberth assumes. The (c) lines give a representation of the φ-domain structure produced in satisfaction of Align-R(XP,φ) theory and assuming strict layering, while the (d) lines give a phrasing produced in satisfaction of Match(XP,φ).

(22) a. u-zile chi-búuku / na méeza '(s)he bought a book / and a table'
b. $_{VP}$[u-zile$_{V}$ $_{NP}$[[chi-búuku]$_{NP}$ [na meezá]$_{NP}$]$_{NP}$]$_{VP}$
c. $_{\varphi}$(u-zile chi-búuku)$_{\varphi}$ $_{\varphi}$(na meezá)$_{\varphi}$ *Align-R(XP,φ)*[33]
d. $_{\varphi}$(u-zile $_{\varphi}$($_{\varphi}$(chi-búuku)$_{\varphi}$ $_{\varphi}$(na meezá)$_{\varphi}$)$_{\varphi}$)$_{\varphi}$ *Match(XP,φ)*[34]

(23) a. mw-ana w-a Núuru / m-someleeló / laazíle
'The child who Nuuru / read to (him) / fell asleep.'
b. [$_{NP}$[mw-ana [w-a $_{clause}$[[Núuru]$_{NP}$ [m-someléelo]$_{VP}$]$_{clause}$]$_{NP}$ [laazíle]$_{VP}$]
c. $_{\varphi}$(mw-ana w-a Núuru)$_{\varphi}$ $_{\varphi}$(m-someléelo)$_{\varphi}$ $_{\varphi}$(laazíle)$_{\varphi}$ *Align-R(XP,φ)*
d. $_{\varphi}$(mw-ana w-a $_{\varphi}$(Núuru)$_{\varphi}$ $_{\varphi}$(m-someléelo)$_{\varphi}$)$_{\varphi}$ $_{\varphi}$(laazíle)$_{\varphi}$ *Match(XP,φ)*

In ChiMwiini vowel length (whether lexical or derived) will surface only in a position at the right edge of a φ, namely in penultimate position, or in ante-penultimate position if the penultimate is light, a pattern familiar from the Latin stress rule. The noun for "child" in (23) has an underlying long vowel, *mw-aana*; its shortening to *mw-ana* in (23) indicates that it is *not* in the R-phrase-edge position for licensing length. As for the phrasal H tone (marked with acute accent here), its default position is on the penultimate syllable in a φ (though certain verb forms and lexical items require H tone on the final syllable of φ). The appearance of phrasal tone and vowel length in a word therefore diagnoses a word's location at the right edge of φ.

Note that in (22c) the R-edge-based parsing proposed by Selkirk (1986) produces a φ-domain structure that is radically non-isomorphic to the syntactic structure: the first conjunct of the conjoined direct object NP is grouped with the preceding verb in a φ. The Match-based parsing, on the other hand, produces the isomorphic φ-domain structure in (22d). Similarly, in (23c), *Nuuru*, the subject of the relative clause is non-isomorphically grouped with the preceding head noun of the NP by Align-R(XP,φ). But the non-isomorphic phrasings are not necessary to the account of vowel length alternations, contra Selkirk (1986). *Mw-aana* loses its underlying vowel length in (23) because it is not in the antepenultimate or penultimate position of a φ, where its length could be licensed by phrasal stress. The vowel shortens in the absence of this positional licensing of vowel length, simply because it is stress-less, not because it is located in the same φ-domain as a following stressed syllable. As (22cd) and (23cd) show, both the Align-R-XP theory and the Match XP theory predict the same locations of right edge of φ, and so both correctly predict the position of phrasal stress and the licensing of vowel length. They also both predict the distribution of the penultimate H tone

accent. It is in fact not possible to decide between these two analyses on empirical grounds, since there is in point of fact no phonological phenomenon that diagnoses the left edge of φ in Chi Mwiini. What drove Selkirk (1986) to adopt non-isomorphic phrasing analyses like those in the (c) lines was the presupposition that the strict layer hypothesis held of any prosodic structure representation, not any evidence that showed that all material between successive right edges was contained in the same φ.

The conclusion that the single-edge-theory-driven φ-domain structure is underdetermined by the facts in ChiMwiini is also true of the facts of Xiamen Chinese reported by Chen (1987). In Xiamen the phenomenon at issue is the distribution of the two possible tonal realizations that any lexical morpheme of Xiamen may display – the phrase-final form and the non-phrase-final form. (There is no known phonological rule that relates the final forms to the non-final forms). Let us assume with Chen that the appearance of the phrase-final form diagnoses the right edge of a phonological phrase (his "tone group"). Chen shows that the right edge of a phonological phrase diagnosed in this way coincides with the right edge of a syntactic XP. But there is no empirical reason to assume with Chen that the φ extends from the right edge of one syntactic XP to the right edge of another. Just as there was no reason to make that assumption in ChiMwiini. The non-final form in Xiamen will appear just as long as the word is not in phrase-final position.

A left-edge version of Align XP was taken by Selkirk and Tateishi (1991) to derive the distribution of major phonological phrases (also known as intermediate phrases known also as maximal φ) that provide the phonological domains for the phonetic implementation of sentence tone in Japanese, as proposed for example in the account of Pierrehumbert and Beckman (1988). The left edge of major phrase/intermediate phrase/maximal φ is the locus of a significant upward pitch reset, which largely undoes the various pitch downtrends that result in a lower pitch range at the end of a preceding phrase. This is illustrated in sentence (24).

(24) a. $[_{NP}[$ [Na^{H*}gano$_N$-no] ani^{H*}yome-ga $]_{NP}$ $_{VP}[_{NP}[$ [Ao^{H*}yama-no]
Nagano-GEN sister-in-law-NOM Aoyama-GEN
yama^{H*}mori-o] yonda $]_{VP}]_S$
mountain guard-ACC called
'A sister-in-law from Nagano called a mountain guard who is in Aoyama.'

b. $_\varphi($^{LH}Na^{H*}↓gano$_N$-no ^{LH}ani^{H*}↓yome-ga$)_\varphi$ $_\varphi($ ^{L}↑^{H}Ao^{H*}↓yama-no ^{LH}yama^{H*}↓ mori-o yonda $)_\varphi$

c. $_\varphi($^{LH}Na^{H*}↓gano$_N$-no ^{LH}ani^{H*}↓yome-ga$)_\varphi$ $_\varphi(_\varphi($ ^{L}↑^{H}Ao^{H*}↓yama-no ^{LH}yama^{H*}↓ mori-o$)_\varphi$ yonda $)_\varphi$

The down arrows in (24bc) indicate the phonetic downstep that is produced in Japanese following a lexically accented syllable, marked with H*, and the up arrow indicates the presence of the significant upward reset attested at the left edge of a major phrase/intermediate phrase/maximal φ. The Align-L XP constraint argued for by Selkirk and Tateishi assigns the prosodic structure in (24b) to sentences like these: the left edge of each XP (the subject and the object and VP edges)

coincides with a φ-edge; the verb is included in the same φ with the preceding direct object. This inclusion of the verb within the same φ as what precedes is consistent with the phonetic facts, given that there is no major upward reset of pitch before the verb. But the Match Phrase analysis which is given in (24c) is also consistent with the known facts. In (24c) the verb *yonda* is external to the φ which parses the object phrase, but is not a φ itself and so would not trigger upward reset. (Note that any φ-structure internal to NP has not been included here.)

In the three cases just reviewed the facts do not decide between a single-edge-based theory and a Match-based theory of syntactic-prosodic constituency, precisely because the reported phonological and phonetic phenomena in these languages diagnose only one edge of φ. The syntactic single-edge that was aligned with the phonological domain in the minimal strictly layered Align XP analysis was precisely the sole edge at which some phonological or phonetic phenomenon was attested. But in a language like Xitsonga, where both edges of φ are diagnosable, it is Match(XP,φ) and Match(φ,XP), which together require a perfect correspondence of left and right edges of XP in syntactic representation with right and left edges of φ in phonological representation, whose predictions hold (cf. Note 30). This can be seen in (25), where the prediction of the Match(XP,φ) and Match(φ,XP) combination is compared with the predictions of Align-R(XP,φ) and of Align-L(XP,φ):

(25) φ-domains predicted by Match vs. Align correspondence for Xitsonga XPs

a. $[\ _{NP}[$ noun adjective$]_{NP}\ _{VP}[$ verb $_{NP}[$ noun adjective $]_{NP}\ ]_{VP}]$

b. Match(XP,φ)/Match(φ,XP): $_{\varphi}$(noun adj)$_{\varphi}$ $_{\varphi}$(verb $_{\varphi}$(noun adj)$_{\varphi}$)$_{\varphi}$

c. Align-R(XP,φ): *$_{\varphi}$(noun adj)$_{\varphi}$ $_{\varphi}$(verb noun adj)$_{\varphi}$

d. Align-L(XP,φ): (i) *$_{\varphi}$(noun adj)$_{\varphi}$ $_{\varphi}$(verb)$_{\varphi}$ $_{\varphi}$(noun adj)$_{\varphi}$

OR: (ii) $_{\varphi}$(noun adj)$_{\varphi}$ $_{\varphi}$(verb $_{\varphi}$(noun adj)$_{\varphi}$)$_{\varphi}$

The Match(XP,φ)/Match(φ,XP) combination straightforwardly predicts the attested recursive, level-skipping domain structure in (25b) with its nested XP corresponding to recursive φ and no φ structure assigned to the verb. Align-R(XP,φ) wrongly predicts that the entire VP should constitute a single φ as in (25c). The prediction of Align-L(XP,φ) is more complicated. If strict layering is assumed, as in the original Selkirk's (1986) proposal, the verb and its following multi-word object would each be parsed as a φ, as in the ungrammatical (25d-i). Yet the attested recursive φ-structure in (25d-ii) is also consistent with Align-L(XP,φ) – even if this correspondence constraint does not force the recursivity. Some other constraint would have to be responsible for the presence of the recursive structure. But, since recursivity as a property of prosodic structure is presumably phonologically marked, no phonological constraint could be given responsibility for producing the prosodic structure in (25d-ii). Moreover, even allowing the option for *both* Align-R(XP,φ) and Align-L(XP,φ) to come into play simultaneously, as proposed by de Lacy (2003), would produce the same pair of possible φ-domain structures, without favoring the recursive one. In other words, a theory of prosodic structure

formation consisting of just the set of single-edge Align-R/L XP constraints on syntactic-prosodic constituency correspondence and a set of phonological markedness constraints cannot derive the pattern of phonological phrasing observed in Xitsonga. In Section 2.3.3 we review the proposal by Truckenbrodt (1999) that adding an additional constraint type – Wrap XP – to the theory of the syntax-prosodic structure relation allows for this recursivity to be produced.

2.3.3 Truckenbrodt (1999): Wrap(XP,φ) and Align R/L(XP,φ) Truckenbrodt (1999) proposes that a theory of syntactic-prosodic constituency correspondence must include, in addition to the constraints Align-R(XP,φ) and Align-L(XP,φ), a constraint Wrap(XP,φ). Wrap XP demands that each syntactic XP be contained in a phonological phrase (φ). Given a VP containing one or more arguments, a single φ corresponding to just the VP will satisfy Wrap XP with respect to that VP as well as to all the component NPs: for example, $_{VP}$[XP XP . . . verb]$_{VP}$ → $_{\varphi}$(XP XP . . . verb)$_{\varphi}$. In this φ-domain structure all XP are *contained within a φ*, as required by Wrap XP. (As shown, Wrap XP on its own does not guarantee the appearance of nested φ structure.) The addition of Wrap XP to a repertoire of constraints on the syntactic-prosodic constituency correspondence that includes Align-R(XP,φ) and Align-L(XP,φ) is designed to provide the foundation for a richer cross-linguistic typology of possible syntasctic structure-φ structure correspondences than is available with Align XP theory itself.

Importantly, the Wrap XP-plus-Align XP theory does permit the generation of recursive φ-structures on the basis of nested XPs in the syntax. This can be illustrated with data from German where a recursive φ-structure is necessary for an adequate account of phrase stress. In German main phrase stress and the pitch accent that necessarily accompanies main phrase stress, falls on the rightmost element within a φ. In the SOV structure found in embedded clauses, discourse-new subject and object phrases each necessarily receive a pitch accent, while the verb does not:

(26) a. . . . weil María die neuen Gesétze studiert
because M. art new law-s study-pres:3s
'because Maria is studying the new laws'
b. . . . weil [[María]$_{DP}$ $_{VP}$[[die neuen Gesétze]$_{DP}$ studiert]$_{VP}$]
c. Match(XP,φ)/
Match(φ,XP): . . . weil $_{\varphi}$(María)$_{\varphi}$ $_{\varphi}$($_{\varphi}$(die neuen Gesétze)$_{\varphi}$ studiert)$_{\varphi}$
d. Wrap XP *and*
Align-R XP: . . . weil $_{\varphi}$(María)$_{\varphi}$ $_{\varphi}$($_{\varphi}$(die neuen Gesétze)$_{\varphi}$ studiert)$_{\varphi}$
e. Wrap XP *and*
Align-L XP: *. . . weil $_{\varphi}$(María)$_{\varphi}$ $_{\varphi}$(die neuen Gesetze studíert)$_{\varphi}$

The Match(XP,φ)/Match(φ,XP) account, coupled with a theory of the prosodic phonology of φ, correctly derives the absence of necessary pitch accent on the verb in all new sentences. These phonological assumptions are (i) that main phrase

stress in φ is assigned to the rightmost ω in φ in German, (ii) that a pitch accent is necessarily assigned to the syllable carrying main stress of φ and (iii) that a φ contains just one main stress. (See Kratzer and Selkirk 2007). By these assumptions, the main phrase stress of the recursive φ structure corresponding to the VP at the right in (26c) is correctly located within the lower φ (on the direct object), where it is marked by a pitch accent.[35] A Wrap XP-plus-Align-R XP can produce the same recursive φ-structure as Match XP and so can also derive the same phrase stress and pitch accenting, seen in (26d): Align-R XP ensures the presence of the right φ edge following the direct object and Wrap XP is responsible for the left and right φ-edges flanking the entire VP, ensuring that the entire VP, including its component constituents, is contained within a same φ.[36] Yet the Wrap XP-plus-Align XP theory of syntactic-prosodic constituency correspondence differs from Match theory in predicting, cross-linguistically, a greater range of possible φ-structures for a given syntactic structure. For example, the combination of Wrap XP and Align-L XP would predict the non-recursive φ-structure in (26e) for a verb phrase with verb and direct object. This structure is inappropriate for German, since it would predict that in all-new sentences like these the verb should necessarily bear main phrase stress and pitch accent and the direct object no phrase stress or pitch accent at all. But Wrap XP-plus-Align XP theory is committed to the existence of this particular VP-φ structure relation in some language.

Cross-linguistic investigation is required in order to ascertain which theory of syntactic-prosodic constituency correspondence provides a better foundation for a typology of φ-domain formation in grammar. The Wrap XP-plus-Align XP theory predicts, of course, that the broader range of φ-domain structures it defines should be attested in some language. Kahnemuyipour 2004, 2009 reports on a highly relevant cross-linguistic investigation of patterns of phrasal stress and pitch accenting according to which a phrase stress pattern like (26e) for a VP consisting of XP plus verb in an all-new sentence is not attested in any of the languages investigated. Match XP predicts this cross-linguistic limitation in distribution of phrase stress prominence, but Wrap XP-plus-Align XP does not. If further research does indeed substantiate Kahnemuyipour's claim that, in systems where the assignment of phrase stress is discernable, phrase stress never falls on the verb instead of the direct object in neutral all-new sentences, then Wrap XP-plus-Align XP theory must be rejected as a theory of syntactic-prosodic constituency correspondence.

According to Truckenbrodt 1999, the full typology of possible XP-φ structure relations that are predicted by the Wrap XP-plus-Align XP theory relies on a role for the prosodic markedness constraint Non-recursivity (Selkirk 1996). If Non-recursivity is higher ranked than Wrap XP and Align-R/L XP, then recursive φ-structures are excluded. In this case, Truckenbrodt (1999) proposes, the respective ranking of Wrap XP and Align XP in the grammar of a language produces a range of non-recursive φ-domain organizations, and these should all be attested in some language(s) of the world. This can be illustrated with the case of a verb phrase containing multiple internal arguments:

(27) φ-domain structures generable from syntactic input [NP NP V]$_{VP}$ by Wrap XP-plus-Align XP theory supplemented by high-ranked Non-recursivity constraint:[37]

a. (NP NP verb)$_{\varphi}$ by NonRec(φ) >> Wrap XP >> Align-R/L XP
b. (NP)$_{\varphi}$ (NP)$_{\varphi}$ verb by NonRec(φ) >> Align-R XP >> Wrap XP:
c. (NP)$_{\varphi}$ (NP verb)$_{\varphi}$ by NonRec(φ) >> Align-L XP >> Wrap XP:

In the case of (27a), where Wrap XP dominates Align XP, the VP itself will correspond to a φ and internal to the VP there will be no further φ-structure. In the case where Align-R XP or Align-L XP dominates Wrap XP, there will be no φ-domain corresponding to the VP itself, but there will be φ-domains marking off the syntactic phrase break between the arguments, as in (27b) and (27c). By contrast, a theory including the correspondence constraints Match(XP,φ)/Match(φ,XP) but lacking any SLH-enforcing markedness constraints like Non-recursivity only allows for the recursive structure in (28):

(28) φ-domain structure generable by Match theory from syntactic input [NP NP V]$_{VP}$: ((NP)$_{\varphi}$ (NP)$_{\varphi}$ verb)$_{\varphi}$

The research question that is now open is whether or not cross-linguistic investigation of φ-domain-sensitive phenomena provides support for the systematic appearance of recursivity envisaged by the Match theory of syntactic-prosodic constituency correspondence articulated above, or whether instead all or some of the richer array of φ-structure possibilities envisaged by the Truckenbrodt (1999) articulation of Wrap XP-plus-Align XP theory are attested.

It must be kept in mind that proposals concerning theories of the syntax-phonological domain structure relation have to be evaluated in the context of a full theory of grammar. As discussed in Section 2.1.3, the phonological component of a grammar includes a theory of language-particular variation in domain-sensitivity as well as a theory of phonological markedness constraints on prosodic structure wellformedness. So explanations for patterns of distribution of domain-sensitive phenomena in the sentences of a language do not rely just on the theory of the syntactic-prosodic constituency relation alone. It is with this general point in mind that we turn to a case which Truckenbrodt (1999) takes to display the non-recursive phrasing of VP illustrated in (27a), that of Tohono 'O'odham, whose sentence phonology was first described and analyzed in Hale and Selkirk (1987).

In Tohono 'O'odham, evidence from the distribution of a basic default (L)HL tonal pattern in the sentence supports the positing of φ-domains over which that (L)HL pattern is defined. Note first that a single (L)HL patterns extends over the clause-final verb and all the XP arguments which precede the verb and lie to the right of the tense-bearing auxiliary on the left, as in (29).

(29) i. L HHH HHH H L
Na-t [[g wákial]$_{XP}$ [g wísilo]$_{XP}$ cépos]$_{VP}$
Inter-Aux:3:perf art cowboy art calf brand:perf
'Did the cowboy brand the calf?'
ii. L H*---------- H*-------- H* L%
Na-t $_{\varphi}$(g wákial g wísilo cépos)$_{\varphi}$

(Orthographic acute accents mark the positions of word stress; the (i) lines contain the tonal transcription given in Hale and Selkirk (1987); the (ii) lines contain the phonological phrasing representations they propose, and a phonological analysis of the tonal patterns to be explained below.) An XP that precedes the auxiliary shows its own (L)HL pattern, as in (30). And any verbal argument that is dislocated to the right shows the (L)HL pattern: compare (31ab).

(30) i. HLL L HHH H L
[g wákial]$_{XP}$ 'at $_{VP}$[[g wísilo]$_{XP}$ cépos]$_{VP}$
art cowboy Aux art calf brand:perf
'The cowboy branded the calf.'
H*LL% L H*-------- H*L%
ii. $_{\varphi}$(wákial)$_{\varphi}$ 'at $_{\varphi}$(g wísilo cépos)$_{\varphi}$

(31) a-i. L HHH HHH HHH H H L
No [[g wákial]$_{XP}$ [g wísilo]$_{XP}$ [g wíjina-kaj]$_{\varphi}$ wúpda]$_{VP}$
Inter-Aux art cowboy art calf art rope-with rope:imperf
'Did the cowboy rope the calf with the rope?'
L H*---------- H*---------- H*---------- H*L%
a-ii. No $_{\varphi}$(g wákial g wísilo g wíjina-kay cépos)$_{\varphi}$
b-i. L H L H LL HLL HL L
No [wúpda]$_{VP}$ [g wákial]$_{XP}$ [g wísilo]$_{XP}$ [g-wíjin-kay]$_{XP}$
'Did the cowboy rope the calf with the rope?'
b-ii. L H* L% H*LL% H*LL% H*L L%
No $_{\varphi}$(wúpda)$_{\varphi}$ $_{\varphi}$(g wákial)$_{\varphi}$ $_{\varphi}$(g wísilo)$_{\varphi}$ $_{\varphi}$(g-wíjin-kay)$_{\varphi}$

Moreover, right dislocation of an XP within a nominal or locative XP results in a sequence (L)HL, as comparison of the patterns in (32ab) and (33ab) shows:

(32) a-i. HH HL
[[g Húsi]$_{YP}$ kíi]$_{XP}$
art Joe house
'Joe's house'
a-ii. H* H*L%
$_{\varphi}$(g Húsi kíi)$_{\varphi}$

b-i. HL HL
[[g kíi]$_{XP}$ [g Húsi]$_{YP}$]$_{XP}$
art house art Joe
'Joe's house'
b-ii. H*L% H*L%
$_{\varphi}$(g kíi)$_{\varphi}$ $_{\varphi}$(g Húsi)$_{\varphi}$

(33) a-i. L HHH HL
[am [míisa]$_{YP}$ wéco]$_{XP}$
loc table under
'under the table'

a-ii. L H*---- H*L%
$_{\varphi}$(am míisa wéco)$_{\varphi}$

b-i. L HL HLL
[[am wéco]$_{XP}$ [g míisa]$_{YP}$]$_{XP}$
loc under art table
'under the table'

b-ii. L H*L% H*L%
$_{\varphi}$(am wéco)$_{\varphi}$ $_{\varphi}$(g míisa)$_{\varphi}$

Because the sequence of (L)HL patterns in cases like (32) and (33) involve XP-internal differences in syntactic phrase structure, the phonological domain for the (L)HL pattern observed here is indeed the φ-domain, which appears in correspondence with syntactic phrases, rather than the ι-domain, which corresponds to syntactic clauses.

The (L)HL pattern associated to each φ-domain may analyzed as follows: a H tone appears on each word stress; a high plateau extends from one-word-stressed syllable to another in the φ (perhaps due to fusion). The syllables following the φ-final word stress bear a L tone. There is also L tone on any syllable(s) preceding the H tone on the first word stress of the φ-domain. It will be assumed that any stress-less syllable that does not come to bear a H tone through spreading is realized with a default L tone. This analysis of default L tone for stress-less syllables is supported by the fact that when the leftmost word of a φ has initial stress the tonal pattern of the φ begins with just a H. By contrast, the fact that a φ-final word-stressed syllable shows a fall from H to L indicates that the final contour must result from the association of the word-stress H tone and a right-phrase-edge L% boundary tone to the same φ-final stressed syllable. The right-edge L% boundary tone is therefore a reliable indicator of right edge of φ. This tonal analysis is reflected in the phonological representations in the (ii) lines above.

Two earlier analyses of the distribution of this (L)HL pattern, and in particular of the absence of the right-edge L% on pre-head XPs, sought to enrich the theory of constraints on the relation between syntactic structure and phonological phrasing in order to accommodate it. Hale and Selkirk (1987) took the distribution of this canonical (L)HL pattern to indicate that – in the grammar of Tohono 'O'odham – the relevant Align-R phrasing constraint for Tohono 'O'odham was parameterized to appeal only to maximal projections that were not "lexically governed." XPs preceding the verb within the VP were lexically governed, as were pre-head XP within nominal and locative phrases. But the pre-Aux XP, lying outside the VP, was not lexically governed; nor was the VP itself. Nor was any XP that was dislocated. This parameterized Align-R(XP$^{\text{not-lex-gov'd}}$,φ) constraint correctly locates the right-φ-edges indicated in (ii) in each of the examples above. Truckenbrodt (1999), on the other hand, opted to restrict the theory of phrasal edge-alignment constraints to just the general type Align XP, and instead offered an account of the Tohono 'O'odham phrasing pattern that relies on two distinct types of universal constraint on the relation between syntactic and prosodic constituency: Align XP and Wrap XP. As discussed above, Truckenbrodt proposes that the ranking of Wrap XP above Align XP in the grammar of a language in the grammar of a

language in which Non-recursivity prevails is the source of the VP-size φ-domain that is observed in Tohono 'O'odham in (29), (30), and (31).[38]

But there is a third possible approach to this data from Tohono 'O'odham within the more restrictive Match theory of syntax-prosodic structure correspondence. Match(XP,φ)/Match(φ,XP) will assign the sentences in (29) and (30) the recursive φ-domain structure corresponding to the VP and its daughter XPs that is seen in (34ii) and (35ii). The pre-Aux XP of (30/35) will be assigned a separate φ.

(34) i. L HHH HHH H L
Na-t [[g wákial]$_{XP}$ [g wísilo]$_{XP}$ cépos]$_{VP}$
Inter-Aux:3:perf art cowboy art calf brand:perf
'Did the cowboy brand the calf?'
L H*---------- H*--------H* L%
ii. Na-t $_{\varphi}$($_{\varphi}$(g wákial)$_{\varphi}$ $_{\varphi}$(g wísilo)$_{\varphi}$ cépos)$_{\varphi}$

(35) i. H LL L HHH H L
[g wákial]$_{XP}$ 'at $_{VP}$[[g wísilo]$_{XP}$ cépos]$_{VP}$
art cowboy Aux art calf brand:perf
'The cowboy branded the calf.'
H*LL% L H*-------- H*L%
ii. $_{\varphi}$(wákial)$_{\varphi}$ 'at $_{\varphi}$($_{\varphi}$(g wísilo)$_{\varphi}$ cépos)$_{\varphi}$

Note that the (L)HL contours found in these sentences are defined over just those φ which are not dominated by any other φ. These are precisely instances of maximal φ, a subtype of φ defined by Itô and Mester (2007, 2010/11). The L% boundary tone can be analyzed as being restricted to the right edge of a maximal φ. This alternative within a Match theory framework to the Hale and Selkirk (1987) and Truckenbrodt (1995, 1999) accounts of the distribution of the LHL phrasal tone pattern in Tohono 'O'odham relies on the theory of domain-sensitivity. We know on the basis of independent evidence that phonological theory must allow for this general sort of language-particular variation in domain-sensitivity when it comes to the distribution of boundary tones at prosodic constituent edges. Japanese shows a L% at the right edge of *any* φ, not just maximal φ; Bengali (Hayes and Lahiri 1991) shows a H% at the edge of any φ. Importantly, a reanalysis of the distribution of phrasal H tone insertion in Kimatuumbi (Odden 1987, 1996) along the lines proposed here for Tohono 'O'odham would locate this H% tone epenthesis at the right edge of a maximal φ.[39]

The hypothesis favored here, then, is that the theory of prosodic domain formation and prosodic domain-sensitivity includes (i) the highly restrictive universal Match theory of syntactic-prosodic constituency correspondence, (ii) a theory of domain-sensitivity in phonology which allows for domain-sensitive phenomena to be sensitive to any of the prosodic category types defined in the theory, in the general manner sketched in Sections 2.1.3 and 2.2.3, and (iii) a phonological theory of markedness constraints on prosodic structure (to be discussed immediately below). Further cross-linguistic research is required, of course, to determine

if this theory does indeed allow for an insightful characterization of typological variation in the distribution of domain-sensitive phenomena in the sentence phonology or phonetics of any language investigated.

3 Phonological Influences on Prosodic Constituent Structure

The existence of non-syntactic influences on phonological domain structure provides the fundamental argument in favor of a prosodic structure theory of phonological domains. In what follows a brief review will be made of the sorts of prosodic markedness constraints that can result in the formation of surface phonological constituents that do not correspond to syntactic constituents, in violation of Match constraints on the syntactic-prosodic structure correspondence. The case study from Lekeitio Basque in Section 3.2 illustrates a range of prosodic markedness effects leading to such instances of non-isomorphism between φ-domains and syntactic phrases.

3.1 Prosodic Markedness Constraints Interacting with Match Constraints

The theory of prosodic structure markedness constraints, which plausibly has its foundations in the purely phonological rhythm-grounded foot, extends its reach to ι, φ, and ω, which, according to the theory of supra-foot prosodic category types put forward above, are grounded in syntactic-prosodic structure correspondence. The vocabulary of phonological constraints includes all prosodic categories, regardless of provenance. We have seen that domain-sensitive phonological markedness constraints – like Non-Finality – mention the syntactically grounded category types ι, φ, and ω. The review below shows that markedness constraints on prosodic structure itself – like BinMin – do so as well.

3.1.1 Size Constraints Constraints requiring that a prosodic constituent be structurally binary at some lower level of prosodic analysis are well motivated at the foot level, where languages divide up according to whether they require feet to be minimally bimoraic or bisyllabic (Hayes 1995 among others). At the ω level prosodic minimality is often a consequence of the fact that a ω consists of at least one foot (which must itself be binary), though Itô and Mester (1992/2003) have shown that certain Japanese loanword adaptations require ω binarity at a higher level, so that the derived ω in such cases must consist of either two feet or a foot plus a syllable. It is to be expected, then, that prosodic minimality effects will be common across languages both at the φ-level and at the ι-level, where, as Ghini (1993) and Inkelas and Zec (1995) suggest, they might override the effects of syntax-prosodic structure interface constraints (see also Selkirk 2000). At higher levels of phonological domain, it has been proposed that syntactic branchingness

affects prosodic domain structure (Nespor and Vogel 1986; Bickmore 1990, among others), though Inkelas and Zec 1995 have argued that such restrictions are not based on syntactic branchingness, but rather on prosodic word (ω) count. In this spirit, Selkirk (2000) proposes the prosodic markedness constraint formulated here as BinMin(φ,ω), requiring that a φ minimally consist of two ω, and the constraint BinMax(φ,ω), requiring that there be no more than two ω in a φ. The ranking of BinMin(φ,ω) with respect to the interface constraint Match(Phrase,φ), discussed above, makes for a clear typological difference between languages. Xitsonga, for example, does not allow single-word noun phrases and ChiMwiini, for example, does:

(36) Typological differences in φ domain structure due to ranking of BinMin(φ,ω):

a. BinMin(φ,ω) >> Match(Phrase,φ) [Xitsonga, Italian, . . .]
b. Match(Phrase,φ) >> BinMin(φ,ω) [ChiMwiini, German, . . .]

As we saw in Xitsonga, the effect of the ranking of BinMin(φ,ω) over Match(Phrase,φ) is to disallow phonological phrases that would correspond to syntactic phrases that are sub-binary. The effect of the excess-size-penalizing BinMax(φ,ω) constraint would be the opposite, if it outranked the input-output correspondence constraint Match(φ,Phrase), which rules against instances of φ in the output representation which do not correspond to some syntactic phrase in the input. Japanese is reported to show a case of this sort: a noun phrase consisting of four lexical words in a recursive left-branching genitive structure has a surface prosodic structure containing a sequence of two binary φ (cf. Selkirk and Tateishi 1988; Kubozono 1993; Shinya, Selkirk, and Kawahara 2004):

(37) Effects of BinMax(φ,ω) on prosodic φ structure in Tokyo Japanese

$[[[\text{N-no N-no}]\ \text{N-no}]\ \text{N-ga}]_{NP} \rightarrow {}_{\varphi}(\ {}_{\varphi}(\text{N-no N-no})_{\varphi}\ {}_{\varphi}(\text{N-no N-ga})_{\varphi}\ )_{\varphi}$

Note that the effect of high-ranked BinMax(φ,ω) is the appearance of a φ – the one embedded on the right – that is not identical to any constituent of the syntactic representation, in violation of Match(φ,Phrase). This optimal prosodic structure departs from the left-branching φ-binarity predicted by Match Phrase, so as to produce an improvement in the binarity of φ-structure.

It is also reported in the literature that there are prosodic size effects on prosodic phrase organization that appear to depend on brute syllable count and are not reducible to prosodic binarity (Delais-Roussarie 1995; Prieto 1997, 2005; Elordieta *et al.* 2003, 2005, D'Imperio *et al*. 2005). The question arises whether such effects give rise to categorical, typological, distinctions between languages, or whether they may reflect more universalist tendencies of performance organization. This is clearly a question for future research.

3.1.2 Left-Edge Strengthening Examination of foot distribution within words testifies to constraints that are specific to prosodic left-edge organization. A class

of languages including English and Garawa, which place main word stress on the rightmost foot, take the option of organizing the left edge of the word into feet, when presented with the choice (Hayes 1995). This so-called "initial dactyl effect" (McCarthy and Prince 1993a) can be seen in five-syllable monomorphemic English words like *Tàtamagóuchi*. If foot organization were organized entirely from right to left, the pattern should be *Tatàmagóuchi*, which contains a "stray" syllable at the left edge of ω. The necessity, instead, of a left-edge foot could be seen as an instance of what will be called a Strong Start effect:[40]

(38) Strong Start

> A prosodic constituent optimally begins with a leftmost daughter constituent which is not lower in the prosodic hierarchy than the constituent that immediately follows:
>
> * (π_n π_{n+1} . . . .

A Strong Start effect at levels of prosodic organization above the word can plausibly be found in avoidance of "stray" syllables or feet at the left edge of phonological phrases, an avoidance seen for example in the promotion of initial weak pronouns to ω status or in their obligatory rightward displacement (Werle 2009; Elfner 2010). A Strong Start effect is also possibly the source of a bias to place a pitch accent on the first prosodic word of an ι-domain in English: avoiding a "stray" ω at the left edge of an ι-domain would involve promoting it to φ status, with consequent φ-stress prominence assignment and the resulting insertion of epenthetic H* pitch accent. In Xitsonga, the parsing of preposed syntactic phrases as ι-domains rather than the expected φ, discussed above in Section 2.1.1, is plausibly another instance of a Strong Start effect. The ι-domain status of preposed noun phrases in Xitsonga constitute a violation of the output-input interface faithfulness constraint Match (ι,Clause); this violation would be produced if the prosodic markedness constraint Strong Start were to outrank Match(ι,Clause) in the grammar of Xitsonga. The grammar of Northern Sotho, which does not show this promotion to ι of preposed XPs, would by contrast rank Match(ι,Clause) above Strong Start.

3.1.3 Prosodic Stress Prominence Assignment Another robust prosodic markedness effect at foot level and above concerns the presence and placement of stress prominence (for which see the classic Hayes 1995 review). A class of constraints calls for a prosodic constituent π to be headed, namely to contain a most prominent, main stressed, constituent (see Selkirk 1980b; 2009 among many others). Call this the constraint family ProsProm(π). Another class of prosodic constraints locates that prominence at the left or right edge of the prosodic constituent π (Prince 1983; McCarthy and Prince 1993a). Call this the constraint family Edgemost-L/R (Prom-π, Edge-π) (Prince and Smolensky 1993; McCarthy 2003). So feet are either trochaic or iambic, with prominence either on the left-hand or right-hand syllable, depending on the language. And main word stress falls either on the leftmost or rightmost foot of the word. The expectation, then, is that at the φ-level, the location

of main stress within φ would be edgemost, on a language-particular basis – falling in the rightmost ω in some languages, the leftmost ω in others. This prediction does appear to be borne out in precisely those cases where it can be put to the test, namely in cases of minimal φ that consist of two ω, as in syntactic phrases consisting of Adjective plus Noun. φ-stress is rightmost in [Adj N] phrases in German, English, Italian and leftmost in Turkish and Persian.

The theory of prosodic stress prominence in the φ-domain is a theory of the default assignment of phrasal stress in sentences which are "neutral" – all new in the discourse. These default stress patterns are claimed to reflect the prosodic constituency of the sentence (see e.g. Nespor and Vogel 1986; Kratzer and Selkirk 2007, 2010). Other approaches to describing default phrase stress patterns have characterized them as depending directly on syntactic constituency (Chomsky and Halle 1968; Selkirk 1984b; Cinque 1993; Kahnemuyipour 2004, 2009; Wagner 2005, 2010; Truckenbrodt 2006[41]), though these could not account for any phrase stresses appearing in constituents that are non-isomorphic with the syntax.

Relevant to the point at hand, the markedness constraints ProsProm(φ) and Edgemost-R/L (Prom-φ, Edge-φ) may have an effect on the very prosodic constituency of the sentence, precisely in cases where the distribution of phrasal stress prominence is not determined by default, and specifically in cases where syntactic constituents are marked for contrastive Focus or discourse-Givenness. It has been proposed that Given-marked constituents in English are submitted to an interface constraint, call it Destress Given, that prohibits them from carrying phrasal stress prominence (see, e.g. Ladd 1980; Reinhart 1995; Féry and Samek-Lodivici 2006; Selkirk 2008 and Kratzer and Selkirk 2007, 2010 The ranking Destress Given >> ProsProm(φ) would lead to an absence of φ-level stress on a Given constituent. This required absence in φ-level prominence for Given-marked constituents would lead to an absence of φ-domain status for a Given-marked phrase, in violation of Match(Phrase,φ), when ProsProm(φ) is higher ranked than Match(Phrase,φ) in the grammar of the language. As for the case of contrastive Focus, it has been proposed by many authors (Jackendoff 1972; Truckenbrodt 1995; Reinhart 1995; Zubizarreta 1998; Szendroi 2001) that a Focus-marked constituent is required to contain the greatest stress prominence within some relevant domain; call this interface constraint Stress Focus. As Truckenbrodt (1995) suggests, the appearance of a φ-domain edge at the right or left edge of a Focus constituent, observed in a variety of languages, could be understood as an effect of the prosodic markedness constraint Edgemost (Prom-φ; R/L; Edge-φ). The phrasal stress that is produced in order to satisfy Stress Focus would induce the presence of a φ-edge adjacent to that stress, through the effect of Edgemost Prom, and could thereby introduce a φ-domain structure that does not correspond to a syntactic phrase (see, e.g. Selkirk 2002, 2009). In cases of this sort, then, there is potential for violation of Match correspondence constraints.

3.1.4 Constraints on The Relation Between Tone and Prosodic Prominence

Another particularly relevant sort of prosodic markedness constraint regulates the relation between tone and prosodic prominence (stress), and by extension the

relation between tone and prosodic constituency. There are languages in which predictable (epenthetic) tone appears on the main stress of the foot (Singapore English: Siraj 2008), on the main stress of a ω (Tohono 'O'dham, see above; Serbo-Croatian: Zec 1999; Werle 2009; Cairene Arabic: Hellmuth 2007; see also Hyman 2006), or on the main stress of a φ (Bengali: Hayes and Lahiri 1991; English: Ladd 1996, 2008a; Selkirk 2000: Féry and Samek-Lodovici 2006; Calhoun 2006; German: Kratzer and Selkirk 2007).[42] (Cases of tone that is restricted in distribution to a local prosodic prominence have standardly been referred to as pitch accents.) It has also been observed that in some languages a lexical tone may migrate to a position of stress prominence, whether in the word or in the phrase, see for example, Kisseberth (1984) on Digo. Yip (2002) hypothesizes that these sorts of phenomena testify to the existence of phonological markedness constraints on the tone-stress relation. Such constraints could be formulated schematically as below:[43]

(39) a. No Toneless π-Stress
The prosodically prominent (stressed) syllable of a prosodic constituent of level π must be associated to some tone T.
b. No π-Stress-less Tone
A tone T must be associated to a prosodically prominent (stressed) syllable of a prosodic constituent of level π.

(In the constraint schemata given here, π is a variable over the set of prosodic category types {foot, ω, φ, ι}.) Markedness constraints on tone-prosodic stress prominence association like those in (39) may – if high enough ranked – contribute to determining the prosodic constituent structure of a sentence, and, in particular, may be responsible for violations of the correspondence constraints that govern the φ-domain/syntactic phrase relations in the sentence. Consider, for example, the well-known fact of Japanese that a syntactic NP with embedded genitive *-no* NP that consists of two accented words [[A-no] A-case] will be prosodically parsed as $_{\varphi}(_{\varphi}($ A-no $)_{\varphi}\,{}_{\varphi}($ A-case $)_{\varphi})_{\varphi}$, that is, into two minimal φ (also referred to as minor phrase or accentual phrases), whereas the same syntactic phrase type with a sequence of two unaccented words will be tend to be parsed as one minimal φ, namely as $_{\varphi}($ U-no U-case $)_{\varphi}$ (see Poser 1984; Kubozono 1988, 1993; Selkirk and Tateishi 1988).[44] In the first case, a lexically accented head noun (the one on the right) acquires φ-phrase status in the phonology. In the second case, an unaccented genitive noun phrase lacks φ-phrase status in the phonology. The hypothesis here is that the two-φ sequence for the two accented nouns, for example, comes about due to a tone-stress markedness constraint (i.e. No φ-Stress-less tone) which requires that each tonal pitch accent be associated with a distinct φ-prominence; this has as a consequence the parsing of the sequence of accented words into two distinct φ.[45] In the following section we will see further examples from Lekeitio Basque of the effect of tone-prosodic prominence markedness constraints on the surface prosodic constituent structure of the sentence.

3.2 A Case Study from Lekeitio Basque

This section reports on the findings of Elordieta (1997, 1998, 2006, 2007a, b, c) and Jun and Elordieta (1997) concerning the φ-domain structure of Lekeitio Basque, a variety of Northern Bizkaian Basque spoken in Spain. These works establish the basic generalizations concerning the patterning of phonological phrase organization in the language, and argue that various prosodic markedness constraints on the composition of phonological phrases outrank syntax-phonology correspondence constraints, producing important cases of non-isomorphism between syntactic and prosodic constituents. The generalizations that Elordieta lays out make a distinction between two distinct types of phonological phrase – the accentual/minor phrase and the intermediate/major phrase. It will be assumed here, with Itô and Mester (2007, 2010/11), that these are both instances of the prosodic category φ. The facts below are consistent with the proposal in Section 2.2.3 that interface Match Phrase constraints appeal only to a single prosodic category φ. At the same time, it will be seen that phonological markedness constraints and rules of phonetic interpretation may recognize distinctions between sub-types of φ that depend on the position of a φ in a recursive φ-domain structure, as the Ito and Mester propose. Recall that a minimal φ is a φ which dominates no other φ, while a maximal φ is a φ that is dominated by no other φ.

Lekeitio Basque is a lexical pitch accent language and its phrasal phonology displays many typological similarities with Tokyo Japanese. There is a contrast between lexically accented and unaccented words:[46]

(40) Lexical contrasts in pitch accenting in Lekeitio Basque:

a.	ama	'mother'	itturri	'fountain'	lagun	'friend'
b.	égi	'truth'	mái	'table'	áurre	'front'

The lexical pitch accent is H*L. At most one pitch accent can appear in a minimal φ, and when it does it must be on the final ω in the φ.[47] These restrictions have the consequence that a minimal φ can consist of a single accented word, one or more unaccented words, or a sequence of an unaccented word followed by an accented word. Two-word noun phrases of these types are given in (41); these are shown with isolation pronunciations, such as might also appear in topic position, for example.[48]

(41) Two-word noun phrase types in Lekeitio Basque:

(i) [[A-gen] A-case] → $_{\varphi}(_{\varphi}(\ A\ )_{\varphi}\ _{\varphi}(\ A\ )_{\varphi})_{\varphi}$

a. [[lagún-en] liburú-ak]
friend-pl.-gen book-pl. abs "the friends' books"

b. LH H*+L !LH H*+L
$_{\varphi}(_{\varphi}$(lagún-en)$_{\varphi}$ $_{\varphi}$(liburú-ak)$_{\varphi})_{\varphi}$

(ii) [[U-gen] A-case] → $_{\varphi}$(U A)$_{\varphi}$
a. [[lagun-en] liburú-ak]
friend-sg. gen book-pl. abs "the friend's books"
b. LH H*+L
$_{\varphi}$($_{\varphi}$(lagun-en liburú-ak)$_{\varphi}$

(iii) [[U-gen] U-case] → $_{\varphi}$(U U)$_{\varphi}$
a. [[lagunen] dirua]
friend-sg. gen money abs "the friend's money"
b. LH
$_{\varphi}$($_{\varphi}$(lagun-en dirua)$_{\varphi}$

As in Japanese, a non-lexical epenthetic LH boundary tone sequence marks the left edge of any φ in Lekeitio Basque, and provides a crucial source of evidence for the φ-domain structures posited. The presence of a LH rise at the left edge of both accented nouns in the (i) case is evidence for its sequential φ-domain structure, in which each accented noun occupies its own minimal φ. That these two accented minimal φ are also grouped together within a superordinate maximal φ is indicated by substantial downstepping of the tone of the second word (marked with "!"); in Lekeitio Basque, as in Japanese, such downstepping would not appear if the second noun were initial in a maximal φ (see below). Of course the presence of the superordinate φ in (41-i) is predicted by the Match(Phrase,φ) interface constraint, and the lower φ corresponding to the genitive NP on the left is too. But, as in the Japanese case discussed just above, the presence of the minimal φ on the head noun on the right must be attributed to a phonological markedness constraint, one whose effect is to allow just one accented noun within a minimal φ. It was suggested above for Japanese that this is a constraint of type (39b): No φ-Stress-less Tone. The imposition of φ-stress by the presence of lexical accent has as a consequence the imposition of the φ constituency implied by the presence of φ-level main stress, and this has as a consequence the epenthesis of the initial left-φ-edge LH tone sequence.

In the (41-ii) case, the absence of a LH rise immediately preceding the accented noun in second position shows that the accented word is not itself a φ, but rather that the UA sequence together forms a single minimal φ (one that is at the same time maximal). The UU case in (iii) also constitutes a minimal φ (that is also maximal). The effect of absence of lexical accent on φ-domain structure in Basque will be discussed below.

In Tokyo Japanese, an accented word triggers a downstepping of the pitch range in which the subsequent word is realized (Poser 1984; Pierrehumbert and Beckman 1988; Kubozono 2007). Elordieta shows that this phonetic effect is found in Lekeitio Basque as well. It is illustrated in the three-accent noun phrase in the sentence in (42), for example (cf. Elordieta 2007c):

(42) a. $_{clause}$[$_{NP}$[$_{NP}$[$_{NP}$[Maialénen]$_{NP}$ lagúnen]$_{NP}$ liburúak]$_{NP}$ gustaten dxákes]$_{clause}$
Maialen-gen friends-gen books-abs like aux
'They like Maialen's friends' books.'
b. $_{\iota}$($_{\varphi}$($_{\varphi}$(Maialénen ↓lagúnen)$_{\varphi}$ ↓liburúak)$_{\varphi}$ ⇓$_{\varphi}$(gustaten dxákes)$_{\varphi}$)$_{\iota}$

(42) shows downstep after every accented word preceding the verb. (For typographical ease and visual clarity, in (42) and other examples that follow, a simple down arrow "↓" indicates the presence of downstep, with the complex down arrow "⇓" indicating the larger-than-normal downstep or pitch compression that is found on the sentence-final verbal complex. An orthographic acute accent replaces the tonal representation of the pitch accent H*L, and the left-edge tonal rise LH will not be represented at all, but should be assumed to be present at the left edge of any φ. Finally, a φ consisting of just a single accented ω will not be written.) A further aspect of the phonetic interpretation of Lekeitio Basque that mirrors that of Tokyo Japanese is the upwards pitch reset that is found at the left edge of a maximal φ. (A maximal φ is dominated by no other φ. "Maximal φ" corresponds roughly to the "major phrase" or "intermediate phrase" in earlier accounts.) This upward reset is indicated with an up arrow "↑," as in (43).

(43) a. $_{clause}[_{NP}[\ _{NP}[$Maialénen$]_{NP}$ lagunári$]_{NP}\ _{NP}[$ liburúak $]_{NP}$ gustaten dxákes$]_{clause}$
Maialen-gen friends-dat books-abs like aux
'Maialen's friends like the books.'

b. $_{\iota}(_{\varphi}($ Maialénen ↓lagunári $)_{\varphi}$ ↑$_{\varphi}$(liburúak $)_{\varphi}$ ⇓$_{\varphi}$(gustaten dxákes $)_{\varphi})_{\iota}$

The preverbal three-noun sequence in (42) constitutes the object noun phrase of the sentence. The three-noun sequence in (43) consists of a two-noun dative object noun phrase followed by a single-word direct object noun phrase. The different patterns of downstep/reset in the two sentence types are a function of differences in maximal φ-domain structure, which in turn mirror the differences in syntactic constituency. Elordieta (2006, 2007c) reports on experimental results showing that the F0 relation between the peaks of the second and third nouns is significantly different in the two cases, with the greater difference in F0 in the case of sentences like (42) attributable to downstep (and lack of upward pitch reset). These same sorts of results have been found for Tokyo Japanese (Selkirk and Tateishi 1991; Ishihara 2008).

Two important ways in which the pitch patterning of Lekeitio Basque sentences differs from that of Tokyo Japanese will be discussed below. Both of these involve cases where the φ-domain constituency of the sentence diverges from that which is predicted by simply matching up φ-domains with syntactic phrases. First of all, as Elordieta (1997, 1998) points out, syntactic constituents consisting of unaccented nouns may fail to correspond to the φ-domain structure that syntactic-prosodic structure correspondence constraints would predict. For example, pronunciations of the sentence in (44), which contains three wholly unaccented noun phrases preceding the final verb sequence, include a rendition with tonal properties justifying the prosodic structure representation in (44b) as well as the rendition in (44c).[49]

(44) a. $[\ _{NP}[$[nire lagunen] alabia$]_{NP}\ [\ _{NP}[$umiari$]_{NP}\ _{NP}[$biberoya$]_{DP}$ emóten] jun da]
my friend-gen. daughter-abs child-dat baby-bottle-abs give-imperf go aux
'My friend's daughter has gone to feed the bottle to the baby.'

LH H*+L
b. $_{\iota}($ $_{\varphi}($ nire lagunen alabia umiari biberoya emóten $)_{\varphi}$ jun da $)_{\iota}$
LH H LH H*+L
c. $_{\iota}($ $_{\varphi}($ nire lagunen alabia $)_{\varphi}$ $_{\varphi}($ umiari biberoya emóten $)_{\varphi}$ jun da $)_{\iota}$

(44b) contains an initial LH rise at the left edge of the sentence, a lexical H*L accent on the verb, and a high plateau extending between them. The absence of any instances of LH rise at the left edge of the noun phrases that intervene indicates that all this material is contained – non-isomorphically – within a single minimal φ. For some speakers, though, the presence of two or more words in the subject noun phrase favors the appearance of a corresponding φ, as seen in (44c), while the remaining single-noun arguments of the verb are (non-isomorphically) grouped with it into a second φ.[50] Such cases show that a purely phonological property like the absence of a lexical pitch accent can have an effect on the establishment of φ-domains and can lead to cases of substantial divergence from the phonological domain structure predicted by interface Match Phrase constraints.

A possible explanation for the violation of Match(Phrase,φ) and Match(φ,Phrase) seen in representations like (44bc), where syntactic phrases lacking lexical pitch accents may fail to get prosodically parsed as φ-domains, would make crucial appeal to a prosodic markedness constraint, as suggested by Elordieta (2007a).[51] The assumption here is that this constraint is No Toneless φ-Stress (cf. (39a)). Lekeitio Basque does not allow pitch accent epenthesis onto the main stress of any φ; Elordieta proposes this follows from a high-ranked constraint requiring faithfulness to lexical pitch accent representations. Given this lack of tonal epenthesis, any lexically unaccented word which were to bear main stress of φ would incur a violation of No Toneless φ-Stress at the surface. But ranking DepT* and No Toneless φ-Stress higher than ProsProm(φ) would mean an unaccented would not bear φ-level stress. This violation of ProsProm(φ) would be minimized or eliminated if the φ itself were forced to be absent, an effect that would be produced by ranking Match Phrase lowest of all. The ranking described would produce the representation in (44b). Of course, the existence of variation in the prosodic structure of sentences like those in (44a), shows that this particular constraint ranking is not the whole story. Elordieta (1998) observes that, as seen in (44c), when an all-unaccented syntactic phrase contains two or more words, some speakers prefer to render the phrase as a φ. A full analysis of the prosodic structure of unaccented phrases is beyond the scope of this chapter and will have to await further research. But these facts do nonetheless testify to the role for non-syntactic factors in determining the phonological domain structure of the sentence in Lekeitio Basque, as Elordieta points out.

A second case which Elordieta offers of phonology-induced non-isomorphism between syntactic constituents and phonological domains in Lekeitio Basque involves sentences in which an initial syntactic constituent is not "heavy" enough, and as a consequence forms part of a φ that includes the following syntactic phrase of the sentence, thereby creating a violation of Match(φ,Phrase). In sentence (45a), for example, Match(Phrase,φ) would predict that each of the single-word noun

phrases found there should have the status of a φ, as in (45b). All these φ would be maximal, that is dominated directly by an ι. Upwards pitch reset is therefore expected at the left edge of each medial noun phrase. But the facts turn out to be different. (45a) is pronounced with a pattern of downstep and upward pitch reset that would be derived on the basis of the prosodic structure in (45c), but not on the basis of the predicted φ-domain structure in (45b).[52]

(45) a. $_{clause}$ [$_{NP}$[Amáiak]$_{NP}$ $_{NP}$[amumári]$_{NP}$ $_{NP}$[liburúa]$_{NP}$ emon dotzo]$_{clause}$
Amaya-erg grandmother-dat book-abs give aux
'Amaya has given the book to the grandmother.'

b. *$_{ι}$($_{φ}$(Amáiak)$_{φ}$ $_{φ}$(↑amumári)$_{φ}$ $_{φ}$(↑liburúa)$_{φ}$ ⇓ emon dotzo)$_{ι}$
subject dative object direct object

c. $_{ι}$($_{φ}$(Amáiak ↓amumári)$_{φ}$ $_{φ}$(↑liburúa)$_{φ}$ ⇓ emon dotzo)$_{ι}$
subject dative object direct object

In the hypothesized prosodic structure in (45c), a maximal φ groups together the subject and dative object; this deprives the dative object of maximal φ status and the upward pitch reset that appears at the left edge of a maximal φ. Instead, the φ-structure in (45c) subjects the dative object to the downstepping found after an accent within a φ. Note that the downstepping pattern seen in (45c) is identical to that which is found with the syntactic structure in (43b), in which the φ that groups together the first two nouns of the sentence does correspond to a syntactic constituent. Clearly, Match Phrase is not responsible for generating the superordinate maximal φ in the case of (45c). This is a case of non-isomorphism between syntactic and prosodic constituency which must have its source in phonological constraints.

A further example where the first syntactic phrase of the sentence is not "heavy" enough is provided by sentences like (46), in which an initial one-word syntactic phrase is followed by a two-word phrase:

(46) a. $_{NP}$[Amumári]$_{NP}$ $_{NP}$[Amáyen liburúa]$_{NP}$ emon dotzat
grandmother-dat Amaya-gen book-abs give aux
'I have given Amaya's book to the grandmother.'

b. *$_{ι}$($_{φ}$(Amumári)$_{φ}$ $_{φ}$(↑Amáyen ↓liburúa)$_{φ}$ ⇓ emon dotzat)$_{ι}$

c. $_{ι}$($_{φ}$(Amumári $_{φ}$(↓Amáyen ↓liburúa)$_{φ}$)$_{φ}$ ⇓ emon dotzat)$_{ι}$

The expected φ-domain structure and consequent pattern of downstep and upward reset is as in (46b). But (46c) shows the actual downstepping pattern attested. Indeed, the experiment results of Elordieta (2006/2007c) show that the downstepping pattern exhibited for (46) is not different from the pattern exhibited for the three-word subject in (42).

In sum, the facts discussed thus far seem to suggest that a prosodic markedness constraint requires that the initial maximal φ within an ι-domain be binary, namely that this φ branch into two ω. Respect for this constraint is proposed in Elordieta (1998) and Gussenhoven (2004). Respect for this constraint would produce prosodic structures like those in (45c) and (46c) in which the initial maximal φ corresponds

to no syntactic constituent in the input. Yet the experimental investigation reported in Elordieta (2006/2007c) shows that mere prosodic word binarity is still not enough: initial syntactic phrases consisting of unaccented noun plus accented noun – namely $_{\varphi}(\ U_{\omega}\ A_{\omega}\ )_{\varphi}$ – are not heavy enough to stand on their own as an ι-domain-initial maximal φ either. Reworded in terms of the Itô and Mester theory of prosodic category types, Elordieta's proposal is that there is a prosodic markedness constraint which requires that an ι-initial maximal φ must branch into two φ,[53] not simply into two ω.

Summing up, Lekeitio Basque illustrates a role for a broad range of phonological constraints which, together in a constraint ranking with syntactic-prosodic constituency correspondence constraints, define the phonological domain structure of a sentence. An ideal of prosodic binarity comes into play in accounting for the last array of facts discussed. The restriction of this binarity constraint to initial position of the ι-domain is in some way reminiscent of the left-edge-specific Strong Start constraints alluded to above. As for the tone-stress markedness constraints that are hypothesized to account for the effects of presence or absence of lexical accent on phonological phrasing patterns, in the analysis suggested, they crucially join with ProsProm markedness constraints that call for any prosodic constituent to carry a main stress or head prominence and faithfulness constraints on the tonal representation. Thus in Lekeitio Basque, markedness constraints on tone, binarity, and stress in prosodic structure all contribute to defining a φ-domain structure that may be at odds with the syntactic structure of the sentence.

3.3 Summary

Evidence has been reviewed here that shows a role for properly phonological constraints as part of a theory of the phonological domain structure of the sentence. It supports the conclusion that influences on the phonological domain structure of a sentence are highly modular; it cannot be accounted for by the theory of syntax alone. Rather, a simple theory of the correspondence between syntactic constituency and prosodic constituency posits a set of universal Match correspondence constraints. These interact in language-particular rankings with phonological constraints of the sort reviewed above to produce a prosodic constituent structure for a sentence which matches up, to greater or lesser degree according to that constraint ranking, with the syntactic constituent structure of the sentence. The defining of the phonological domain structure of a sentence is in this sense a true syntax-phonology interface phenomenon, with contributions from the theory of syntactic representation, the theory of phonological representation, and the theory of the correspondence relation between the two.

ACKNOWLEDGMENTS

Research for this chapter was provided in part by NSF grant #BCS-0004038 to Elisabeth Selkirk. I have benefitted from input from innumerable colleagues and students over the

years, and only most recently and directly by interaction with Laura Downing and Annie Rialland in the context of their grant BANTUPSYN (funded through the CNRS-ARN/DFG French/German collaboration), and especially with Gorka Elordieta. I regret that the full list is too long to include; the chapter already exceeds reasonable space limits.

NOTES

1 For useful reviews on issues related to the spellout of morphosyntactic features, see Embick and Noyer (2007), Elordieta (2007d), and Wolf (2008). Phonological properties associated with the information structure features marking focus, givenness, topic and the like are considered here to be cases of morphosyntactic feature spellout and are not examined in this chapter, except in passing (see Section 3.3).

2 Much of this latter area is the traditional domain of the field of syntax. A role for phonological factors in determining some aspects of word order, including those that involve the distribution of focus, has been advocated by Inkelas and Zec (1990), Reinhart (1995), Zubizarreta (1998), Szendroi (2001), Arregi (2002), Samek-Lodovici (2005), Richards (2010), among others.

3 The theory of the syntax-phonology interface reviewed in Truckenbrodt (2006) forms part of the set of "mixed theories," in the sense that it countenances both an independent prosodic structure over which phonological and phonetic phenomena are defined, and a direct appeal to syntactic constituency representation on the part of the phrase-stress-assigning principle Stress XP. An alternative view, of course, is that phrase stress is assigned only indirectly on the basis of syntax, on the prosodic phrasal constituent domains that are themselves defined with respect to syntax.

4 Selkirk (1974) reports that only consonants forming part of an inflectional ending make liaison between a word that is head of a phrase and a vowel-initial word at the beginning of the phrasal complement that follows. In pre-head contexts, liaison is not so restricted. Pak and Friesner (2006) presents data showing that the surface prosodic constituents revealed by intonational patterns in French cannot provide the context for liaison and argues instead that liaison is introduced in the process of morphosyntactic spellout, independent of prosodic domain (see also Pak 2008).

5 This useful term is due to Itô and Mester (2010/11). A level skipping configuration constitutes a violation of the phonological constraint Exhaustivity in the Selkirk (1996a) proposal decomposing the stricter layer hypothesis into a number of distinct constraints on prosodic domination. Of these, Exhaustivity and Nonrecursivity are violable.

6 For example, the notion that recursivity is a systematic property of prosodic domain structures–contra the strict layer hypothesis – has been emerging with particular force in recent years (see Ladd 1986 *et seq*.; Selkirk 1996; Frota 2000; Dobashi 2003; Féry and Truckenbrodt 2005; Wagner 2005, 2010; Itô and Mester 2007, 2010/11). Section 2.4 addresses the significance of these findings for current theories of the interface.

7 Section 2.3 provides a reanalysis of certain cases that have been assumed to show nonisomorphism that is brought on by satisfying syntactic-prosodic constituency correspondence constraints.

8 Section 2.2.2 presents Match theory as a type of correspondence theory in the sense of McCarthy and Prince (1995), and distinguishes two versions of correspondence constraint, one requiring that a designated syntactic constituent have a corresponding prosodic constituent in phonological representation and another requiring that a surface prosodic constituent correspond to a constituent in syntactic representation.

9 Proposals in the context of minimalist phase theory include Dobashi 2003, Revithiadou and Spyropoulos (2005, 2009), Adger (2006), Selkirk (2006), Ishihara (2007), Kratzer and Selkirk (2007), among others).

10 The term domain that is used in this chapter refers to an abstract constituent structure that controls the phonological and phonetic interpretation of the sentence. It is not identified with any particular aspect of tonal or segmental representation. The notion of domain introduced in Kisseberth (1994) and developed in Cassimjee and Kisseberth (1998) concerns the representation of tone, and is designed, in part, to supplant the autosegmental representation of tone and tonal spreading. This chapter draws on the generalizations about the relation between tonal domains (= tonal spreading) and constituency-related domains that have been brought to light in Kisseberth (1994).

11 It is plausible that penultimate lengthening is in fact a reflex of penultimate stress prominence assigned on the ι-domain, in which case the phenomenon that is ι-domain-sensitive would be stress assignment, not lengthening. This issue cannot be decided here.

12 The transcription of Xitsonga examples is from Kisseberth (1994), as are the translations. The illustrative glosses which accompany certain of the examples have been supplied by one of the editors of this volume.

13 The exclamation point in the examples in (6) and below indicates that the high tone on the following syllable is downstepped. Downstep appears when a high tone is preceded by another, distinct, high tone in the same domain. In (6ab), the H of the HL sequence on *â* triggers downstep of what follows; in (6cd) the downstepping in *tí-ho:m!ú* is due to the H tone that spread onto the first syllable of the word from the preceding verb.

14 See Downing (2011) on an alternative approach to explaining the asymmetry between right and left dislocation structures in Bantu.

15 Kisseberth (K153) reports that high tone spread from the first word to the second within a noun phrase is also blocked from spreading onto the final syllable of the noun phrase, for example, $[$ xoná xi-ambalo$]_{NP}$ > (xoná xí-ámbálo)$_{\varphi}$. This is expected, since a multi-word noun phrase will always correspond to a φ, and so Nonfinality(φ,H) will do the blocking at the right edge.

16 The formulation of H-Spread here expresses the marked status of a configuration in which a H tone associated to a syllable to the left fails to spread onto a toneless syllable on the right (cf. Myers 1999).

17 This maintenance of final lexical H is predicted if Nonfinality(φ,H) is ranked lower than the anti-deletion faithfulness constraint MaxTone (cf. McCarthy and Prince 1995). As for the permissibility of high tone spread onto lexically toneless syllables, it implies the ranking of H-Spread above whatever faithfulness constraint that disallows tone-syllable associations that are not part of underlying representation.

18 Note that this shows that H-Spread must dominate the constraint Nonfinality(ω,H), which holds at the level of prosodic word. See Section 2.1.3.

19 It is not the presence of other, lexical, tones in the object noun phrase in (13ii) that explains the lack of high tone spread. The OCP does block high tone from spreading to a syllable that is adjacent to a lexical high tone but the (c) example shows that H tone can in principle spread from the verb into a following single-noun object that has lexical tone, as long as one syllable intervenes between the two H.

a. ndzi-vóná xí-xlámbétwá:na 'I see a cooking pot.'
b. ndzi-vóná ma-k!ó:ti 'I see vultures.'
c. ndzi-vóná vá-la:l!á 'I see enemies.'

20 There is evidence that high tone does not spread from a subject into a following verb phrase, even when the verb phrase consists of just a single verb. The examples all involve first- or second-person subject pronouns, since only these can be followed by a subject agreement prefix on the verb that is toneless, and thus capable in principle of showing the effects of high tone spread from the subject, for example, *hiná h-a-hle:ka* 'as for us, we are laughing' (K153). The fact is that the final H tone on the pronoun *hiná* does not spread onto the verb; this shows that the verb phrase must be preceded by a left φ-edge. And this also shows that in the case of verb phrase the binarity constraints no φ-domain seen with noun phrase are not observed: a VP will correspond to a φ regardless of whether it contains more than word or not. No attempt is made here to account for lack of binarity effect.

21 Other formal characterizations of this type of edge-sensitive constraint would be required if a Kisseberth-inspired tonal-domain-based representation of H tone spreading span were assumed. In any case, what is to be ruled out is a configuration where a H tone feature spreads across, or a H tone domain includes, the edge of a constituent domain (cf. Note 9).

22 Section 2.2.1 articulates a Match theory as a theory of constituent faithfulness and expresses the Match constraints as correspondence constraints (McCarthy and Prince 1995). The formulation Match(Phrase,φ) given here is a syntactic-prosodic structure correspondence constraint calling for any phrase in syntactic representation (the input) to have a corresponding φ in phonological representation (the output). It is not violated by an output φ which does not have an correspondent in the input syntactic representation.

23 This theory of phonological phrasing makes the typological prediction the ranking of Match(Phrase,φ) and BinMin(φ,ω) might be reversed in the grammar of some other language, in which case, all phrases would be parsed as φ, regardless of their internal word count. Among the Bantu languages, ChiMwi:ni, to be discussed below in Section 2.3, is a language of this sort.

24 It's conceivable that Constraint X here is the prosodic markedness constraint Strong Start, cf. Section 3.1.2.

25 These tableaux only indicate input-output violations of Match correspondence constraints, though output-input violations are in general relevant too. See 2.2.1.

26 We still have to contend with the fact mentioned in Note 18 that a single-word VP will be parsed as a φ, in violation of BinMin(φ,ω). A possible solution would lie in distinguishing more than one type of Match Phrase constraint, with the one relevant to VP ranked above BinMin(φ,ω). See relevant discussion in Section 2.2.2.

27 In addition, the marking of syntactic constituents for information structure properties like contrastive focus, discourse-givenness, and topic-hood may also, whether directly or indirectly, have an influence on the prosodic phrasing structure of a language. See Section 3.3 for a brief treatment of this question, and Lee *et al*. (2007) for a collection of papers documenting such effects.

28 Evidence that CrispEdgeL(ι,H) is active in Xitsonga comes from the set of left dislocation examples in (7). A lexical final H tone does not spread from a preposed NP onto a toneless subject or other preposed NP that follows. Since these following phrases contain just a single word, they do not count as φ, and so it cannot be CrispEdgeL(φ,H) that is blocking H-Spread here. Rather, the blocking is due to the left edge of the ι-domain that follows the lexical final H, more specifically to the ranking of CrispEdgeL(ι,H) over H-Spread.

29 The OCP is another family of constraints that should be expected to show language-particular ranking with respect to H-Spread. As Myers (1997) has shown with evidence

from Bantu, two H tones in sequence constitute an OCP violation only when they are associated to adjacent syllables. Data from Xitsonga shows that the notion 'adjacent syllable' must be relativized to prosodic domains. Kisseberth shows that H-Spread may spread to the final syllable of a verb even if the following single-word direct object noun begins with a lexical H tone. According to the present analysis, these H tones belong to different ω. So ω-internal syllable adjacency is permitted. But as we saw in Note 17, H-Spread does not allow spreading from a verb into a following noun and onto a syllable adjacent to lexical tone further to the right in the noun. This would create a ω-internal configuration consisting of two adjacent H-toned syllables. Defining a set of OCP constraints specific to the distinct prosodic category types and allowing various rankings of H-Spread amongst them predicts a typology of OCP adjacency effects across languages. Xitsonga must have the ranking OCP(ω,H) >> H-Spread >> OCP(φ,H).

30 In phase theory (Chomsky 2001), the TnsP that is complement to Comp constitutes the Spell-Out domain of the CP phase.

31 A recent formulation of a Lexical Category Constraint that accomplishes this is in Truckenbrodt (1999, 2006).

32 (The ν head introduces the subject argument in its Specifier position.)

33 The representation of φ-domain structure in (22c) is the *minimal* strictly layered φ-domain analysis that is consistent with Align-R(XP,φ), not the only one. Also consistent with Align-R(XP,φ) would be φ-domain structure in which the verb stands on its own as a φ, as in the ungrammatical (22e), which would wrongly predict presence of a pitch accent and possibility of realization of underlying vowel length in the verb.

(22e) $*_{\varphi}(\text{ verb })_{\varphi}\,{}_{\varphi}(\text{ NP })_{\varphi}\,{}_{\varphi}(\text{ conj NP })_{\varphi}$

In an optimality theoretic account, some additional constraint – yet to be determined– would be required to rule out this non-optimal non-minimal candidate.

34 It is in fact the combination of the S-P correspondence constraint Match(XP,φ) and the P-S correspondence constraint Match(φ,XP) which predicts the φ-domain structure in (22d). Match(XP,φ) alone would allow for the parsing of the verb as a φ, as in (22f):

(22f) $*_{\varphi}(\,{}_{\varphi}(\text{ verb })_{\varphi}\,{}_{\varphi}(\text{ NP })_{\varphi}\,{}_{\varphi}(\text{ conj NP })_{\varphi}\,)_{\varphi}$

But assignment of φ-domain status to the verb, which lacks XP status here, is ruled out by Match(φ,XP), which requires that any φ in the surface phonological representation correspond to an XP in syntactic constituent structure.

35 Féry (to appear) proposes a Match XP account of these same cases in German, with the same assumptions about the prosodic phonology of stress and pitch accenting. Kratzer and Selkirk (2007), building on the Kahnemuyipour (2004) phase-based theory of German stress, propose a version of Match theory which derives the desired prosodic phonology (φ-domains, main φ-stress and pitch accenting) in (28c), as well as that on intransitive verbs in all-new sentences in German in function of their position in the Spell-Out domain of a phasal head.

There is a certain variability in the accenting of the verb in all-new sentences in German. A slightly less common verb, for example, *untersucht* 'investigates' in the same context might show an accent. In such a case, like the preceding XP arguments, the verb would carry the φ-level stress that gets a pitch accent and would have the

status of a φ. A violation of Match(φ,XP) is brought about in such cases. For this violation to come about, whatever constraint it is that calls for this optional prominence on the verb would have to be higher ranked than Match(φ,XP). An interesting question for future research is just what the nature of that constraint would be.

36 Note that the Wrap XP/Align-R XP combination is satisfied by either the level-skipping recursive φ-structure of (26d), where the verb is not a φ itself, or by a sequential φ-structure within the higher φ that contains the VP in which the verb and object XP are also both parsed as a φ. The latter, non-minimal, recursive structure would have to be ruled out by some additional constraint cf. Note 31.

37 In all cases, these are the *minimal* φ-domain structures that satisfy the constraints at issue, namely Wrap XP, Align XP and Non-recursivity, cf. footnote 31.

38 Truckenbrodt proposes that, in general, the higher node produced as a consequence of an adjunction operation, for example, by right dislocation of YP, as in $[\ XP\ YP]_{XP}$, is not visible to syntactic-prosodic correspondence constraints like Wrap XP. This assumption explains why the internal φ-domains appear the prosodic structure of the dislocated examples in (32b) and (33b), for example. The same assumption will be made for Match Phrase.

39 Clearly, some phonological markedness constraint calling for the presence of a tone at a prosodic constituent edge of the appropriate level must be responsible for the epenthesis of phrasal edge tones in cases like these. See (39) for analogous markedness constraints governing the relation between tone and prosodic stress prominence.

40 McCarthy and Prince 1993a propose an alignment of PWd and Ft to account for the initial dactyl effect, but this does not generalize to the cases below.

41 Truckenbrodt (2006) proposes a constraint Stress XP whose role is to account for the presence of main phrase stress within syntactic phrases. But Stress XP is redundant in a theory of grammar that also posits a syntactic-prosodic constituency correspondence constraint like Match(XP,φ) and a set of prosodic structure markedness constraints like ProsProm(π) that call for prosodic constituents in general, and in particular φ, to contain a prosodic stress prominence.

42 This epenthesis can be seen as a phonological enhancement of abstract prosodic stress prominence (see Smith 2002 on positional markedness).

43 These constraints can be seen as a generalization of the original autosegmental well-formedness constraints (i) "A tone-bearing unit must be associated with some tone" and (ii) "A tone must be associated to some tone-bearing unit" (Goldsmith 1976), on the assumption that tone-bearing units correspond to the class of prosodically defined prominences, ranging from mora *tout court* to mora that is the head prominence of a prosodic phrase.

44 The maintenance in surface forms of the distinction between lexically accented and unaccented words implies, of course, that a faithfulness constraint against epenthesis of tone must outrank the markedness constraint No Toneless π-stress, which would call for the epenthesis of tonal accent in the lexically unaccented case.

45 Alternatively, the limitation of one pitch accent per phrase could be the result of a constraint simply stipulating that a minimal φ (= minor phrase or accentual phrase) contain at most one pitch accent (cf. Selkirk 2000; Gussenhoven 2004). But it is more interesting, theoretically, to attempt to explain such facts in the context of a general autosegmental theory of the relation between tones and tone-bearing units. The notion tone-bearing-unit (cf. Goldsmith 1976) is generalized here in being based on local prosodic prominence, whether the tone-bearing-unit is defined (within the syllable) as a moraic segment, or on higher domains as a moraic segment which bears in addition

some higher level of prosodic prominence. See Hellmuth (2007) for a development of this idea.

46 In contrast to Tokyo Japanese, in Lekeitio Basque this contrast appears only in singular forms; plural nouns are always accented. Also, in Lekeitio Basque the lexical pitch accent always appears on the penultimate syllable in the word, regardless of the morpheme of origin in the word. Tokyo Japanese verbs and adjectives whose roots are lexically accented show this penultimate positioning of accent, but nouns do not.

47 These generalizations are expressed by Elordieta (1997 *et seq*.) using the term "accentual phrase" to identify the relevant prosodic constituent type. The term "minimal φ" used here refers to the same prosodic constituent.

48 All examples here are either from the cited Elordieta (1997 *et seq*.), or from Elordieta directly (personal communication).

49 Experimental studies of comparable Japanese sentences which contain a sequence of wholly unaccented noun phrases have not shown a tendency for unaccented arguments of the verb to join into a same φ (Selkirk, Shinya, and Sugahara 2003; Selkirk, Shinya, and Kawahara 2004).

50 Examples like (44c) reveal the presence of a H boundary tone appearing at the right edge of a φ. This predictable boundary H is not observed when a word carrying H*L pitch accent ends the φ, in which case the L is associated to the φ-final syllable. Epenthesis of a boundary H is avoided in that case, presumably to avoid the creation of a contour tone and still maintain the L.

51 Elordieta (2007a) suggests a role for a constraint with the effect of No Toneless φ-stress in his optimality theoretic account of the resistance of unaccented words to the bearing of the prosodic prominence associated with contrastive Focus.

52 Comparable facts are not reported for Tokyo Japanese, see for example, Pierrehumbert and Beckman (1988), Selkirk and Tateishi (1991), Kubozono (1993), Ishihara (2008).

53 Elordieta sees this as a type of positional markedness constraint, see Smith (2002, 2005). (His own formulation is an intonational-phrase-initial intermediate phrase must dominate two distinct accentual phrases.)

15 Intonation

MARY E. BECKMAN AND
JENNIFER J. VENDITTI

1 The Scope of This Chapter

As a technical term in phonological descriptions of spoken languages, *intonation* refers to patterned variation in voiced source pitch that serves to contrast and to organize words and larger utterances. In this general statement of its meaning, it is synonymous with the technical term *tone*. In typical usage, however, the two terms are differentiated by applying them to different aspects of these linguistic uses of pitch, a differentiation that is reflected in this edition of the *Handbook of Phonological Theory* by the fact of there being a separate chapter on tone (Hyman, this volume). In order to delimit the scope of the present chapter,[1] therefore, we begin by listing the aspects of the linguistic use of pitch that are typically invoked in differentiating intonation from tone.

The differentiation is exemplified by the two parts of the sixth definition for the entry for "tone" in the *Concise Oxford English Dictionary* (eleventh edition): "(in some languages, such as Chinese) a particular pitch pattern on a syllable used to make semantic distinctions" and "(in some languages, such as English) intonation on a word or phrase used to add functional meaning." This sixth definition is tagged as the meanings for a technical term in phonetics, and its second part subsumes the term "intonation," which is defined in its own entry as "the rise and fall of the voice in speaking." In the *COED* entry, then, the primary sense of tone as a technical term in describing sound patterns refers to a localized melodic event (a note or glissando) occurring over the span of a syllable, whereas tone *qua* intonation refers to a pattern of glissandi distributed over a longer span. Also, tone in this primary sense invokes a system of contrastive pitch patterns that act as minimal word-differentiating elements, comparable to the

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

inventory of vowels or consonants of a language, whereas tone *qua* intonation invokes other functions, such as mirroring the syntactic structure of an utterance or indicating its pragmatic role in the larger discourse context. These two sets of contrasting characteristics make for a multidimensional taxonomy of phonetic form in relationship to linguistic function. This much is uncontroversial.

A third aspect of the *COED* definition is more controversial. The two parts of the definition refer to two different sets of languages, reflecting the claim in many broad-stroke surveys such as Hyman (2006b) that particular values along the dimensions of form and function tend to coincide in ways that are conducive to a one-dimensional classification of language types, with "some languages, such as Chinese," at one end and "some languages, such as English," at the other. Careful descriptions of specific languages at every point along the purported continuum, on the other hand, typically use the terms together in ways that defy the typology. For example, in many descriptions of specific Chinese dialects, such as Chang (1958), "tone" is used to refer to the localized melodic events (notes or glissandi) that contrast one-syllable words in citation-form utterances, but "intonation" is also used: to designate notes and glissandi that occur at phrase edges (rather than on designated syllables) with functions other than that of lexical contrast (e.g. marking interrogative speech acts), and to refer to longer-term modulations of the implicit melodic scale that defines the relationship of contrast among different notes and between two different rising or two different falling glissandi. Conversely, in many accounts of the English intonation system, such as Halliday (1967), Goldsmith (1978), and Pierrehumbert (1980), "tone" is used to refer to glissandi or notes that are localized to linguistically significant positions, such as the stressed syllables of some words and the edges of phrases.

The reference to constituents such as "phrases" and "stressed syllables" in these descriptions invokes another technical term – *prosody* – which also occurs in many definitions of intonation, and vice versa. In these definitions, the two words are often treated as being in a hypernym-hyponym relationship whereby "prosody" is the broader cover term that groups various aspects of the pitch pattern of an utterance together with a motley group of other phenomena that defy the assumption of "segmental idealization" – that is the assumption that "speech . . . can appropriately be idealized as a string of ordered discrete sound segments of unspecified duration" (Ladd, this volume). For example, the *COED* defines "prosody" first as "the patterns of rhythm and sound used in poetry" and by extension as "the patterns of stress and intonation in a language." In order to fully delimit the scope of this chapter, therefore, we need to lay out our assumptions about what prosody is, both in relationship to the pitch patterns of a language and in relationship to the rhythms of its vowel and consonant patterns.

In the rest of this chapter, then, we will elaborate on the various aspects of intonation that are invoked in these definitions of tone and of prosody. We begin by briefly laying out one current understanding of what prosody is, and by illustrating this understanding in terms of how one particular language's word rhythms can be described using the representational device that is adopted in much current research on intonational phonology (Section 2). We then review the

various ways in which phonologists and phoneticians have represented the pitch patterns that they observe in the laboratory and the field (Section 3). We go on to describe how research in phonology and phonetics over the decades since the development of such crucial analytic tools as the source-filter theory of vowel production (Fant 1960) has contributed to the evolution of a taxonomy of the forms (Section 4) and functions (Section 5) of spoken language melody. In these descriptions we will use many examples of how the taxonomy applies to varieties of Chinese and of Germanic (including English) in order to emphasize why we think it is not as useful to delimit intonation from tone in terms of sets of languages, a point to which we will return in our summary (Section 6).

2 Defining prosody

2.1 Bases for a Definition

We define prosody as the set of syntagmatic relationships that hold among an utterance's tone, vowel, and consonant specifications, as distinct from the paradigmatic relationships of similarity and contrast that hold between these specifications and any other tone, vowel, and consonant specifications that might stand in closely analogous syntagmatic relationships to each other in some other utterance. We suspect that in the course of acquiring a spoken language, these syntagmatic relationships become reified as an abstract structure projected away from the phonetic substance of the paradigmatic relationships, and that this projection then serves as "a rhythmic scaffolding that specifies designated temporal points of convergence and structural alignment among different components of the grammar" (Arbisi-Kelm and Beckman 2009: 109). The primary device that we adopt to describe this projected syntagmatic structure is the "prosodic tree" – a term first used by Nespor and Vogel (1979) to refer to a particular type of directed acyclic graph that satisfies a set of substantive constraints on the types of node that can occur at different distances from the root and on relationships of (temporal) order among nodes at the same distance (see Selkirk 1980b and Chapter 6 of Pierrehumbert and Beckman 1988, among others).

The language that we will use to illustrate this device is the Tokyo dialect of Japanese, a language that has played a critical role in the development of computationally explicit compositional theories of the elements of intonation contours and their relationship to the prosodic organization of utterances. For example, Fujisaki and Sudo (1971) and Pierrehumbert and Beckman (1988) provide two different accounts of Japanese intonation patterns in which phrase-level pitch range is specified by continuous phonological control parameters that are independent of the categorical specification of more local tone shapes in the lexicon. This incorporation of continuous (or "paralinguistic") specifications of phrasal pitch range directly into the phonological description was a critical element in the development of what Ladd (1996/2008a) calls the "Autosegmental-Metrical" (AM) framework.

Ladd coined this term to refer to a class of models of intonational phonology that began to develop rapidly in the late 1970s and early 1980s, especially after Pierrehumbert's (1980) incorporation of Bruce's (1977) seminal insights about the composition of Swedish tone patterns into a grammar for English intonation contours. In the decades since, the AM framework has been widely adopted in descriptions of intonation systems and in comparisons of prosodic organization across a variety of languages. (See Jun 2005 for a recent collection of such descriptions and Gussenhoven 2004 for a review of work in this framework since Ladd 1996.) All AM models are deeply compositional, analyzing sound patterns into different types of elements at several different levels. The most fundamental level is the separation of material properties from the structural positions that license them. This is roughly comparable to the separation of lexis from phrase structure in a description of the rest of the language's grammar.

2.2 *Autosegmental Content*

More specifically, the "A" (or autosegmental) part of an AM description refers to the specification of content features that are autonomously segmented – that is, that project as strings of discrete categories specified on independent tiers rather than being bundled together at different positions in the word-forms that they contrast (see Goldsmith 1976a, 1979; Hyman, this volume). For example, in Japanese, changing specifications of stricture degree for different articulators can define as many as six different manner autosegments within a disyllabic word-form, as in the word *sanpun* [sampuɴ] 'three minutes', where there is a sequence of stricture gestures for sibilant airflow, for vowel resonances, for nasal airflow, for a complete stopping of air, and then a vowel and a nasal again. However, there are constraints on which oral articulators can be involved in making these different airflow conditions, so that only one place feature can be specified for the middle two stricture specifications in a word of this shape. Therefore, in the grammar of Japanese (as in many other languages), place features can be projected onto a different autosegmental tier from manner features. This projection allows for a very general description of the allomorphic alternation among labial nasal [m] in *sanpun*, dental nasal [n] in *sandan* 'three steps', velar nasal [ŋ] in *sangai* 'third floor', and uvular nasal [ɴ] in the citation form *san* 'three'. This kind of place assimilation for coda nasals applies without exception in the derivational and inflectional morphology of the language, and it leads to differences between first-language speakers of Japanese and Korean in their ability to identify place of articulation of coda nasals in English (Aoyama 2003).

A key insight of Firth (1948), which was developed and elaborated as the basis of autosegmental phonology by North American interpreters such as Goldsmith (1976a, 1979) and Haraguchi (1977) is that the consonant and vowel "units" that are named by symbols such the [s], [a], [m], [p], [u], and [ɴ] in *sanpun* are not necessarily "the basic constituent of a linguistic system" (see Pike 1943a: 42, as discussed in Ladd, this volume). Indeed, several long-standing phonological puzzles, such as "the non-uniqueness of phonemic solutions" (Chao 1934) disappear

when consonant and vowel segments are treated not as an automatic reflection of a "natural phonetic segmentation" but as an epiphenomenon, a by-product of the intersection of place and manner specifications for the word-form when the array of parallel streams of autosegments on different tiers is collapsed into a one-dimensional string. Thus, the "autosegmental insight" (Hyman, this volume: Chapter 7) gradually led to a key insight of CV phonology (Clements and Keyser 1983) and dependency phonology (Anderson 1987): the idea that "vowel" and "consonant" are not merely cover terms for two broad sets of autosegmental substance, but also the names (typically abbreviated as V and C) of two basic types of syntagmatic relationship, which are analogous to the relationships of head versus adjunct in the syntactic organization of words. These two relationships can be represented as two types of leaf node in the prosodic tree that depicts the "M" part of an AM description of a spoken utterance.

2.3 *Metrical Frames*

This "M" (or metrical) part an AM description, then, specifies the hierarchy of prosodic constituents that is the meter (rhythmic scaffolding) of an utterance containing the word-form. For example, in Japanese, there is a low-level prosodic constituent syllable (typically abbreviated σ) that coordinates the place and manner features of words to alternate between more sonorant V manners that are licensed to occur at the head and the less sonorant C manners that are licensed to occur only at the edges. In an utterance of *sanpun*, the syllables dominate a CV level meter that is CVC.CVC. In many other words, there is a more regular rhythmic alternation between the V and C leaf-node types, as illustrated by the two utterances shown in Figure 15.1. All but six of the 26 syllables in these utterances dominate a perfectly alternating CV sequence.

The exceptional cases are of two types. The first is the bare V at the beginning of the verbs *oyoideru* and *oborete* and of the auxiliary verb *-iru* in utterance (b) in the figure. These three V-shape syllables, like the 20 CV-shape syllables, are short. That is, each contains just one mora (abbreviated μ), a constituent between the syllable and the CV leaf nodes that licenses the specification of vowel place and height features at the V nodes within a syllable. The other three exceptional syllables are long. A long syllable contains at least one extra adjunct mora after the head V that is the sole obligatory leaf constituent. This following mora can be either a V or a C. If it is a V, it can be the second part of a geminate vowel, as in the particle *yoo* at the end of utterance (b), or it can be a less sonorous vowel than the head vowel, as in the second syllable of *oyoideru* in the same utterance, where the [i] has lesser sonority than the preceding head vowel [o]. If the following mora is a C, it can be a moraic nasal, as in each of the two syllables of *sanpun* discussed earlier, or it can be the first part of a geminate obstruent, as in the second syllable of the verb *hashitteru* of utterance (a) in Figure 15.1.

As *sanpun* shows, a long syllable that is closed by the moraic nasal can be word final. By contrast, a long syllable that is closed by an obstruent must have a following syllable to provide the second C position for the necessarily geminate

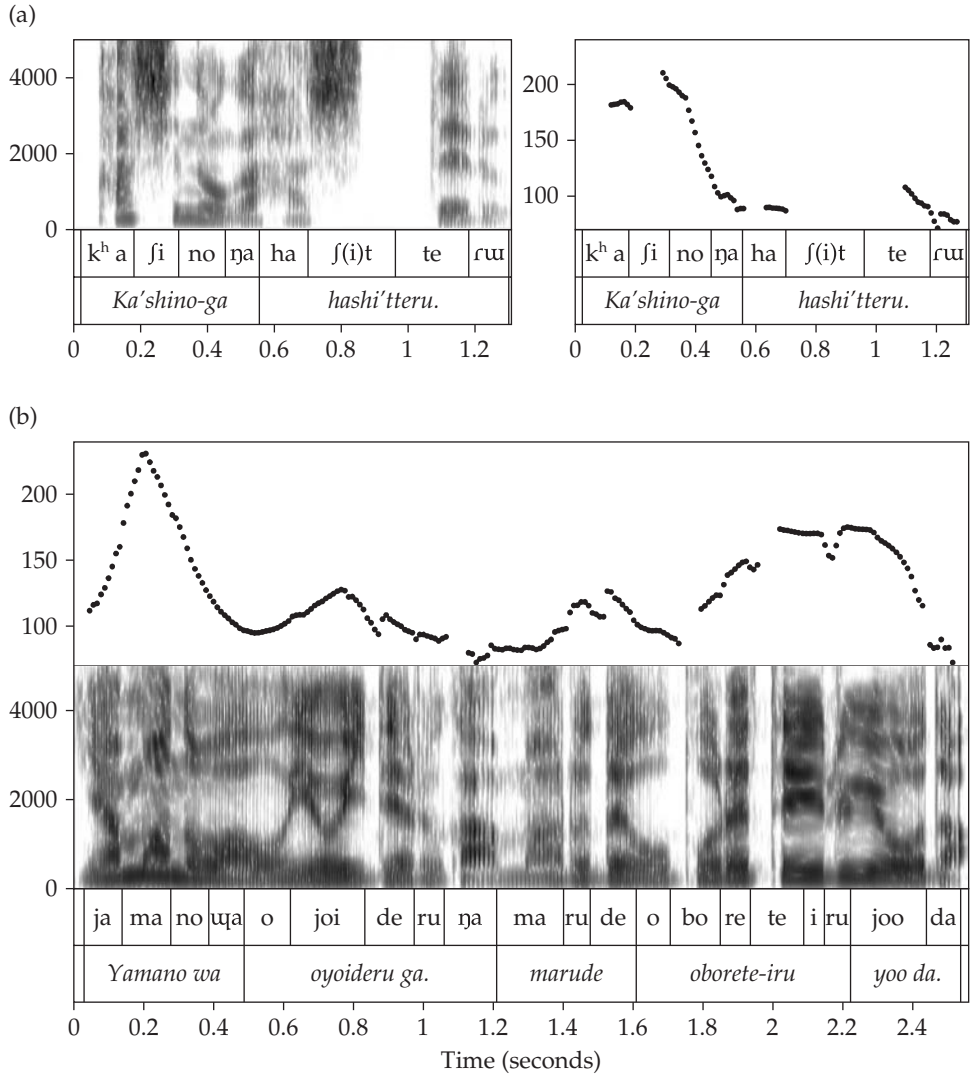


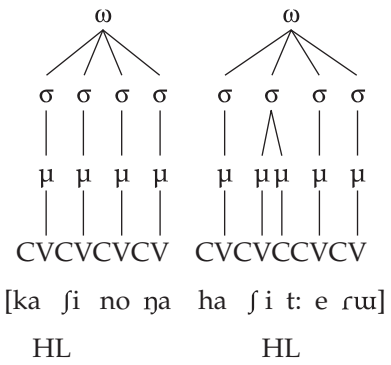
Figure 15.1 Spectrograms and f_0 contours for utterances of (a) *Kashino-ga hashitteru.* 'Kashino is running.' and (b) *Yamano-wa oyoideru ga, marude oborete-iru yoo da.* 'Yamano *is* swimming, but he's nearly drowning right now.' (The utterances are from Venditti, Maekawa, and Beckman 2008.)

consonant. This constraint is one of the motivations for positing prosodic constituents above the syllable as well as below. That is, the place and manner autosegments for a geminate consonant necessarily associate to a sequence of C nodes that is medial to a prosodic word (abbreviated ω).

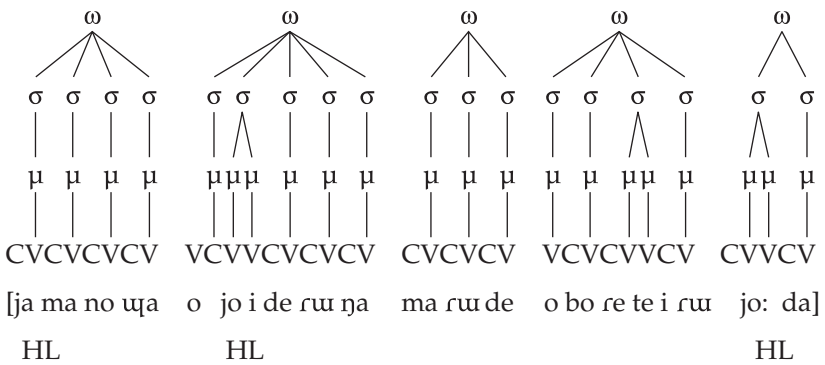
In (1), we show a prosodic tree representation of the ω and ω-dominated constituents for each of the utterances that we have represented phonetically in the

three panels of Figure 15.1. Below each tree we transcribe the sequence of consonant and vowel segments that results when the array of place and manner autosegments is collapsed into a one-dimensional string. The spectrograms in Figure 15.1 provide one kind of (partial) phonetic representation for these place and manner autosegments. Below the transcription of the consonant and vowel segments, we transcribe the 10 tonal autosegments that are specified to occur around the C and V constituents of lexically designated syllables in five of the words that occur in the two utterances. The f_0 contours in Figure 15.1 provide one kind of (partial) phonetic representation for these 10 autosegments and several other tonal autosegments that we will discuss at more length in Section 4.1.

(1) a.



b.



2.4 Prosodic Trees

First, though, we want to bring out more clearly the relationship between the metrical frames that are represented symbolically in (1) and the autosegmental content features that are represented parametrically in Figure 15.1. To do this, we return to the analogy that drives the choice of the representational device in (1). The hierarchy of prosodic constituents in the phonological description of an utterance is analogous to the hierarchy of syntactic constituents in its syntactic description. By this analogy, the prosodic constraint that each syllable must dominate a

V (vowel) category phone that is positioned to be its head is like the syntactic constraint that each verb phrase must have a V (verb) category word positioned to be its head. It is in this sense that we say that the inventory of segments for a language is analogous to the lexis of the language.

One argument in favor of this analogy is that the prosodic category of a phone such as [i] in Japanese is inherently ambiguous in the same way that the syntactic category of a word such as *dam* in English is ambiguous. In different morpho-syntactic contexts, *dam* can be parsed either as the verb meaning "to build an obstruction across a stream or river" or as the noun meaning the resulting obstruction. Similarly, the palatal constriction that gives rise to the low first formant and high second formant on either side of the medial [o] in the word *oyoideru* in spectrogram (b) in Figure 15.1 can be parsed either in terms of a C constituent that is directly dominated by σ (where it is traditionally transcribed as [j]) or in terms of a V constituent that is dominated by the adjunct second μ (where it is traditionally transcribed as [i]).

A second argument in favor of this analogy is the way that empty structural positions are interpreted. That is, the implicit but unrealized subject noun phrase in the second clause of utterance (b) in Figure 15.1 can be recovered from the syntactic parse of the rest of the clause. Similarly, the implicit head vowel in the second syllable of *hashitteru* in utterance (a) can be recovered from the prosodic parse of the rest of the ω even though there is no interval containing the usual acoustic traces of the vowel manner specification in the signal. This token of the segment is "devoiced" and so cannot show the vowel formant values that cue the explicit [i] in *oyoideru*. Nonetheless, the listener parses the presence of some implicit head V constituent for the second σ of *hashitteru* from the fact that [ʃt:] is not a grammatical sequence of phones for any position in a prosodic word of the language. Moreover, the listener parses an associated sequence of high (H) and low (L) tonal autosegments here, even though there is no periodicity to give rise to a sense of pitch on this lexically "accented" syllable, because the melody over this whole sentence is "underlyingly" (i.e. prosodically) identical to the melody of the *Yamano-ga oyoideru* clause in utterance (b) in this figure.

A good grasp of how listeners resolve prosodic ambiguities and of how they recover the intended autosegmental features of an implicit prosodic node in phonological ellipsis is critical for understanding the ways in which prosodic structures above the word are manipulated to fill the functions that we associate with intonation as distinct from tone. In introducing the prosodic constituents C, V, μ, σ, and ω in this section, we have defined them primarily in terms of the distribution of place and manner autosegments for consonant and vowel phones. In describing what listeners do to parse the prosodic categories of ambiguous segments or to recover elided segments, we have made passing reference to the autosegemental content of these structures, using the spectrograms in Figure 15.1 as a convenient representational device for the place and manner autosegments. The two other panels in the figure show a convenient representational device for the 10 tonal autosegments that differentiate words such as *Kashino* and *oyoideru* from words such as *marude*, a representational device that we discuss next.

3 The Representation of Tone and Intonation

3.1 Phonetic Representations

Because of the well-behaved psychophysical relationship between the percept of pitch and the fundamental frequency (f_0) of the periodic voice source, the choice among phonetic measures that one could use to represent tone and intonation patterns is fairly straightforward. Fry (1968) describes several early methods for estimating f_0 values over shorter or longer stretches of speech. These included measuring the durations of successive periods identified in oscillographic records (as in Denes and Milton-Williams 1962) and tracing the frequency of some higher harmonic visible in narrow-band spectrograms (as in Lehiste and Ivić 1963). More recently, the ready availability of computer programs for estimating f_0 using autocorrelation-based algorithms in free signal-analysis packages such as Praat (Boersma 1993) and WaveSurfer (Sjölander and Beskow 2000) has made the f_0 contour of a recorded utterance an even more obvious choice as a first-pass phonetic representation of its intonation or tone pattern. The obviousness of this choice is reflected in the name by which the f_0 contour often is called. In both Praat and WaveSurfer, the f_0 estimates that are returned by the autocorrelation algorithm are called "pitch" values, and many phonologists and phoneticians use the phrase "pitch track" in referring to a time plot of a sequence of estimated f_0 values over some interval of recorded speech.

It is important to remember, however, that the f_0 contour is only a very rough first-pass phonetic representation of the melodic pattern of an utterance, for at least three reasons. The first is that f_0 is not reliably estimated in stretches of speech where less regular source qualities such as creaky voice are in play. The failure of standard algorithms for estimating f_0 in such regions makes the "pitch track" a poor phonetic representation for melodic events that harness such a "non-modal" voice quality as a cue, as illustrated in Figure 15.2. The f_0 contours in the middle panels of the figure were calculated over utterances of two sentences of Putonghua (PRC Standard Chinese) produced by an adult female speaker from Songyuan City in Jilin Province. The creaky voice in the third syllable of each utterance is a cue to the very low pitch target that characterizes this tone, as shown by Gårding *et al*. (1986), among others.

The second reason is essentially the same as the first, but applies to regions where the f_0 is well-defined. When we see the concentrations of energy in a 600–800 Hz band at 0.27 seconds in each of the spectrograms in Figure 15.2, we can read this setting of the filter resonances in terms of the combined labial and dorsovelar constriction gestures for the initial [w] of the surname *Wáng* and for its devoiced allophone after the initial [x] in the surname *Huáng*. By contrast, when we see the f_0 value of 190 Hz at both 0.13 and 0.41 seconds in the right-hand utterance (i.e. at points midway through the [a] nucleus in the two vowels before and after the [xw] of *Huáng*), we cannot know what combination of pulmonary and laryngeal gestures produced this setting of the glottal source period. There

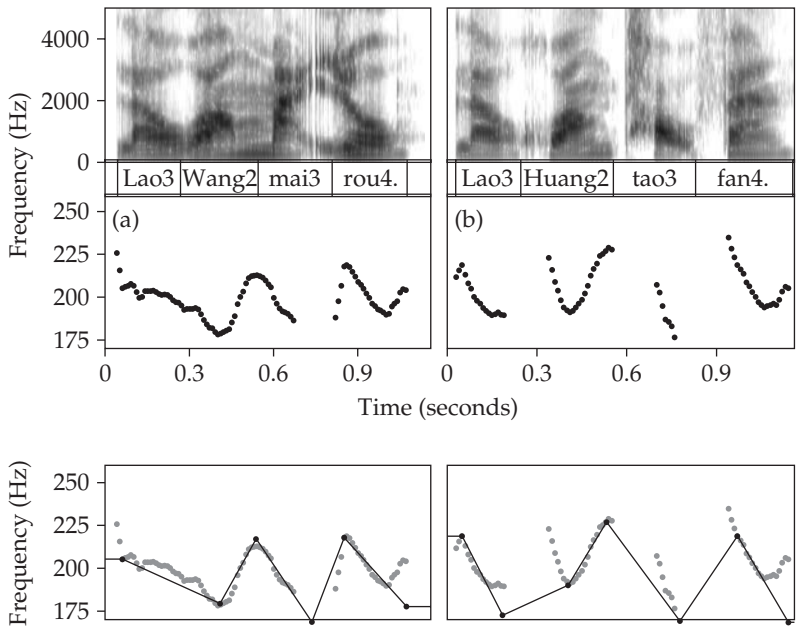


Figure 15.2 Spectrograms and f_0 contours for utterances of (a) *Lǎo Wáng mǎi ròu.* [laʊ^{21}.waŋ^{35}.mai^{21}.z̩ʊ^{51}] 'Old Wang buys meat.' and (b) *Lǎo Huáng tǎo fàn.* [laʊ^{21}.xwaŋ^{35}.taʊ^{21}.fan^{51}] 'Old Huang begs for food.' with a close-copy stylization (solid line) overlaid on each f_0 contour in the bottom row of panels.

is a perceptibly lower tone target for the vowel at 0.13 seconds as compared to the target for the vowel at 0.41 seconds, and this percept of different targets matches transcriptions using Chao's (1930) tone letters. Typically, the tone in *Lǎo* in this phrase-medial context is transcribed as a drop or as a dipping down to the bottom of the tonal space (i.e. [laʊ^{21}]~[laʊ^{214}]), whereas the tone in *Wáng* is transcribed as a rise from a middle region (i.e. [waŋ^{35}]).[2] Further support for positing such a difference in the target minima for these two tones comes from electromyographic studies (Sagart *et al.* 1986), which show consistently high activity of the sternohyoid for the very low-pitched target in the low or dipping tone of *Lǎo* but less reliable involvement of this muscle at the beginning of the rising tone of *Huáng*. The percept of a lower minimum target for [laʊ^{21}] than for [xwaŋ^{35}] in Figure 15.2 suggests that we hear the speaker's intent to produce the creaky voice quality that cues the lower tone target even when the glottal source wave is regular enough that the f_0 tracking algorithm does not fail.

The third reason that the "pitch track" can only be a first-pass phonetic representation of the melody of an utterance is the existence of so-called *micro-prosodic* effects, whereby the aerodynamics of producing contrastive properties of consonants can cause systematic variation in f_0 on consonants and neighboring vowels that is related to the percept of the consonants rather than to the percept of the

utterance tune. These effects also are illustrated in Figure 15.2. The syntactic structure and the sequence of lexical tones are identical between the two sentences. The substantial differences in f_0 shape for the second, third, and fourth syllables are due to the different manners of articulation for the syllable-initial consonants. When native speakers listen to utterances such as these, they parse these micro-prosodic effects for what they are, and perceive the intended tone sequence that is common to both despite the marked differences in f_0 shape (see, e.g. Reinholt Peterson 1986; Silverman 1986).

These three limitations of the f_0 contour can be overcome to some extent by the use of analysis-by-synthesis techniques such as the "close-copy stylization" method pioneered by researchers at the Institute for Perception Research (IPO) in Eindhoven (see, e.g. Cohen and 't Hart 1967; 't Hart and Collier 1975; de Pijper, 1983). A *close-copy stylization* is defined as a synthetic approximation to the melody of the utterance which meets two criteria: "it should be perceptually indistinguishable from the original, and it should contain the minimum number of straight-line segments with which this perceptual equality can be achieved" (Nooteboom 1997: 646). This kind of downsampling of the f_0 contour is very easy to do today, because of the re-synthesis utilities based on LPC analysis or PSOLA (Atal and Hanauer 1971; Moulines and Charpentier 1990; also see Carlson and Granström 1999) that have been implemented in many free signal-analysis packages. The bottom row of panels in Figure 15.2, for example, shows a close-copy stylization of each of the f_0 contours in the figure, created using the implementation of PSOLA re-synthesis in Praat (Boersma and Weenink 2007).

3.2 *Analysis by Synthesis*

Making such a close-copy stylization is a first step in developing and testing a phonological representation of the intonation pattern of an utterance in the models of the British English and Dutch intonation systems described in de Pijper (1983) and 't Hart, Collier, and Cohen (1990). Both of these models pick out some of the line segments in a close-copy stylization as corresponding to phonologically significant events. Two types of phonologically significant event are identified: prominence-lending movements that are anchored to stressed syllables, and juncture-marking movements that occur at the edges of phrases. The phonological significance of a line segment is determined by the criterion of "perceptual equivalence" (as opposed to the "perceptual equality" of the close-copy stylization). Two stylized contours are equivalent if listeners perceive them to have the same utterance melody. In this framework for modeling utterance melody in languages such as English and Dutch, cataloguing the recurring patterns of sequences of line segments in "perceptually equivalent" melodies is analogous to cataloguing the inventory of contrasting vowels or consonants in transcriptions of words and phrases elicited in the field. That is, determining the patterns of melodic equivalence and dissimilarity across a sufficiently large and varied corpus of utterances in several iterations of analysis and perceptual testing of the re-synthesized utterances should yield the set of "melodically distinct pitch movements" for the

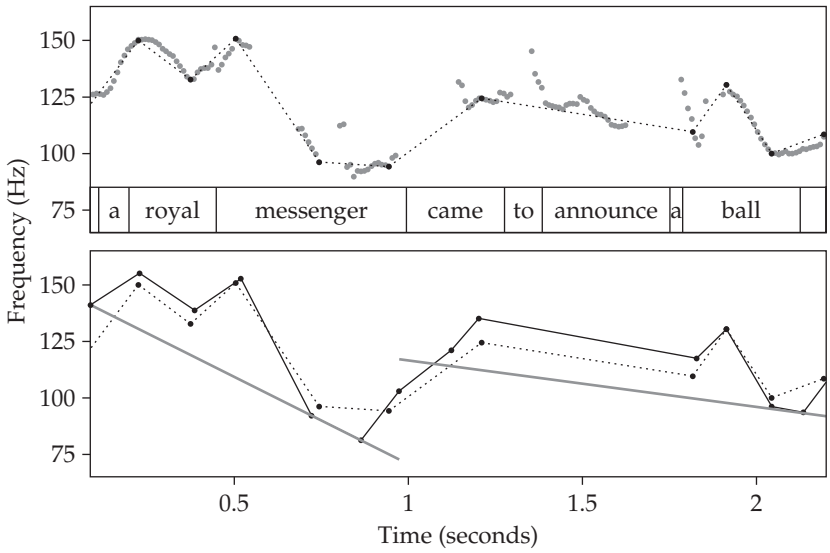


Figure 15.3 Extract from a reading of the *Cinders* passage in the IViE corpus (Grabe 2004) with a two-accent "hat pattern" on each syntactic phrase. Dotted lines are a close-copy stylization overlaid on the original f_0 (top) and on approximations to de Pijper's (1983) model of the melodic elements in this pattern (bottom).

language variety. The model parameters that contrast these line segments, then, are homologous to other distinctive feature sets for the language, such as its contrasting vowel heights or frication source places.

In de Pijper's model of British English, for example, there are eight melodically distinct glissandi which are parameterized in terms of their direction (rise versus fall), their slope (steep versus shallow), and the pitch levels between which they move (e.g. a half rise from lower to middle differs from a full rise from lower to upper levels). Some of these melodic elements are illustrated in the two panels of Figure 15.3. The top panel shows the original f_0 contour and a close-copy stylization of a two-phrase utterance produced by a young male speaker of British English. The lower panel repeats the close-copy stylization and overlays an approximation to the re-synthesized contour that would be generated by de Pijper's model. The intonation pattern in each phrase is the variant of the "hat pattern" depicted in Figure 5.6b in de Pijper (1983). There is a steep prominence-lending half rise early in each phrase (on the first syllable of *royal* in the first phrase, and on *came* in the second) followed by a steep prominence-lending full fall near its end (on the first syllable of *messenger* in the first phrase and on *ball* in the second). The only other phonologically significant line segments are the two "continuation rises" over the last part of *messenger* just before the inter-phrase boundary and over the second half of *ball*.

In addition to these parameters that specify local rises and falls, there are two other essential components of the model. One is the parameter set specifying the

timing of each rise or fall relative to the segments of the syllable or at the phrase edge that licenses it. In de Pijper's model of English, for example, an early steep fall that is prominence-lending starts early enough to be completed at the onset of the vowel in the prominent syllable, a neutral fall starts 30 ms after the vowel onset, and a late fall does not start until after the end of the syllable.

The other essential component of the model is the "declination line" that is the implicit lower level for the melodically distinct rises and falls and that describes the f_0 over sections in between the melodically distinct movements. In complex utterances such as the two-phrase extract in Figure 15.3, there will be as many declination lines as there are phrases, with "reset" at the phrase boundary. The local declination line is steeper for shorter phrases and shallower for longer ones, as suggested by the two overlaid gray lines in the figure.

3.3 Phonological Representations

We have presented de Pijper's model in detail because this kind of analysis by synthesis has proved to be an invaluable tool for going from a database of phonetic representations of a good sample of utterances to an adequately formalized system of phonological representation for languages such as English, which do not offer the fieldworker the crutch of lexical contrast that supports the indispensable initial methodology for studying utterance melody in languages such as Beijing Mandarin or Hakha Lai (see Hyman, this volume: Chapter 7, Section 2). Ladd (2008: 13) lauds this "IPO theory of intonational stucture" as "in many ways the first . . . serious attempt to combine an abstract phonological level of description with a detailed account of the phonetic realization of the phonological elements." Other formal frameworks that laboratory phonologists began to develop at about this same time also used an analysis-by-synthesis approach to decompose f_0 contours into contributions from three model components for (i) the set of localized pitch events, (ii) the timing of these events relative to landmarks such as vowel onsets in prosodically relevant syllables, and (iii) aspects of backdrop pitch range such as the reset points and declination slopes in the IPO model.

Of course, different frameworks make different claims about the allocation of responsibility among these three components, as well as different claims about the appropriate set of parameters internal to each one. However, all fully formalized frameworks have this kind of compositional phonetics, so that the configuration of parameters of the synthesis model that generates a f_0 contour that is perceptually equivalent to each original f_0 contour for a set of utterances that share an intonation pattern can be construed as the phonological representation of that pattern. Formalizing the phonological representation of tune in terms of analysis-by-synthesis model parameters in this way gives phonologists a way to compare models of a language across frameworks, or to compare models of different languages within a framework.

One point of comparison is the size of the smallest sequential element assumed. Where the IPO framework takes melodically equivalent glissandi to be the basic atomic units, many other models of English and Dutch decompose each rise or

fall into a finer-grained sequence of endpoint notes. Pierrehumbert (1980) and Gussenhoven (1984), for example, both analyze the steep prominence-lending fall of the English "hat pattern" in Figure 15.3 as a transition from a higher pitch target to a lower pitch target. These targets, then, were called "high tone" (abbreviated H) and "low tone" (abbreviated L) in an explicit analogy to the use of the terms in work such as Benedict's (1948) description of Thai and Cantonese, and Ward's (1948) description of Efik and Igbo. The analogy made it possible to draw directly on the same kind of compositional phonetics that was beginning to be applied to languages such as Igbo in order to understand better the interplay between the lexically-specified melodic elements that are the "tonemes" of the language and any melodic elements that are produced or parsed "post-lexically" for sequences of words or phrases in their larger discourse contexts. The seminal example of this kind of model is Bruce's (1977) description of Stockholm Swedish word tones, which inspired critical elements of Pierrehumbert's (1980) model of American English intonation patterns as well as of Pierrehumbert and Beckman's (1988) model of Tokyo Japanese word-accent patterns in sentential context.

The analogy worked in the other direction, too, making it clear, for example, that the segmentation grain for utterance melodies is a theoretically interesting question even for languages with lexically contrastive tone. Thus, where Ward (1948) and others analyze Igbo word and phrase patterns as sequences of tone levels, with high or low specified for each syllable, Clark (1978) proposes a system of "dynamic tones" so that a rise or fall is specified only at linguistically significant syllable junctures, as in the IPO-framework models described above. Similarly, where Kindaichi (1957) and Haraguchi (1977) analyze Japanese word- and phrase-level melodies in terms of a succession of low or high tones specified on all moraic segments, Kawakami (1957), Hattori (1961), and Fujisaki and Sudo (1971) analyze them as combinations of underlying rises and falls. We will return to this point in Section 4.1 after describing the ramifications of such differences among models for the symbol sets that are a more typical referent of the term "phonological representation."

3.4 Symbolic Representations

The control parameters for producing and parsing spectral patterns of utterances are often referenced symbolically using transcription systems such as the International Phonetic Alphabet (IPA), in which each basic segmental unit is represented by a letter symbol. The second syllable of the utterance displayed in the upper-left panel of Figure 15.2, for example, can be transcribed as a sequence of three segments [w], [a], and [ŋ], whereas the second syllable of the paired utterance to its right might be transcribed as these three segments plus an initial [x]. While phonologists disagree on the ontological status of such symbolic tags (see Ladd, this volume) there is an overwhelming consensus that this grain of syntagmatic discretization is a useful starting point for phonological analysis, and hence, that the IPA provides a useful common vocabulary for annotating recordings made in the laboratory or the field and for comparing models of what talkers and listeners

implicitly know about how to differentiate words such as the surnames Wáng and Huáng in these sentences.

There is no comparable standard alphabet for segmenting and tagging pitch patterns. The 1989 Kiel revision of the IPA resolved a rivalry between Africanist and Sinological conventions for tagging pitch patterns in languages such as Igbo versus Cantonese, but only by including both transcription systems. In his summary of discussion by the Working Group on Labeling of Suprasegmentals at the Kiel meeting, Bruce (1989: 36–37) describes the failure to achieve even such a laissez-faire resolution for tagging pitch patterns in languages such as English:

> There exists an apparent need for a direct way of symbolizing intonation in a phonetic transcription. However, the opinions diverge regarding the exact way of transcribing intonation. For a phonological transcription of intonation the symbolization is very much dependent on the language and the analysis.

Why might tone and intonation contrasts be so much less amenable to a "phonetic" transcription than are consonants and vowels? We think this is because there is no natural universal segmentation for the pitch pattern shorter than the utterance as a whole. That is, there is nothing akin to the segmentation of filter-resonance patterns afforded by the abrupt transition from a stop-like closure into a more open vocal tract in the CV-like "frame" of canonical babbling (MacNeilage and Davis 2000) and in the "vocal motor schemes" (McCune and Vihman 1987) that become the infant's first words. Across cultures, mothers may use a common stock of attention-getting tunes to draw very young infants into sessions of imitative turn-taking, as suggested by Papoušek, Papoušek, and Symmes (1991). And this common stock of pre-verbal melodies may be a basis for trends such as the prevalence of raised pitch and rising terminals in yes-no questions, as noted by Lieberman (1967), Bolinger (1978), and Ohala (1983a), among others. However, these are not universals of syntagmatic alternation within the utterance. They do not provide a compelling "natural phonetic segmentation" of the melody into units that are any smaller than the tunes of the alternating interlocutor turns.

This difference in pre-verbal rhythmic base gives the symbolic transcription of linguistically significant pitch variation a fundamentally different status from the symbolic transcription of linguistically significant spectral variation. Alphabetic transcription can be a useful pre-theoretical tool for identifying events that are very likely to have phonological significance in the vocal tract filter pattern. The analogous use of symbolic transcription for voice source events makes sense only within the context of a research community in which shared expectations about tunes and their relationship to prosodic constituents above the level of CV units can emerge. For example, the Africanist transcription system that is included in the IPA evolved in a community of Bantuists and of researchers working on non-Bantu languages in the West African Sprachbund, where a long history of language contact has given rise to striking commonalities in syntactic and in prosodic organization above the word. The IPO symbols for English and Dutch

prominence-lending shapes and boundary pitch movements, similarly, evolved in a community of researchers who developed the symbols as shorthand for particular sets of parametric specifications in a shared analysis-by-synthesis model of the intonation systems of these two closely related language varieties.

3.5 Parametric Representations in Prosodic Phonology

The lack of a universal pre-verbal rhythmic base for segmenting speech melodies at any level below the whole utterance may also explain why languages with lexical tone contrasts figured so prominently in the early development of a long lineage of frameworks that led to the emergence of the AM framework. That is, lexical contrast typically provides a more compelling functional basis than pragmatic contrast for segmenting the melodic contour into units smaller than the whole utterance, and there is a long history of applying ideas developed for the description of prosodic structures in languages with lexical tone contrasts to languages such as English. An early example is Kingdon's (1939) application of the notion of "tonetic stress" in the development of a model of English intonation contours that was the first to recognize that stressed syllables before the syllable bearing the "nucleus tone" also can be marked by tones. More recent examples include Liberman (1975), Leben (1976), and Goldsmith (1978), as well as Pierrehumbert (1980) and Gussenhoven (1984), all adapting notational devices from descriptions of lexical tone patterns in various African languages to the description of English utterance tunes.

In reading earlier work in this lineage, we are often struck by the congruence between concepts that were assumed in describing, for example, "the tone-phrasing system of Kongo" (Carter 1974) and the parameterization of intonation patterns in analysis-by-synthesis models such as de Pijper's (1983) model of British English described earlier. In particular, whether lexically-specified tones are transcribed by diacritics on the vowels (as in the Africanist tradition) or by numerals after each syllable (as in the Sinological tradition), utterance melodies are typically described in terms of a convolution of two parts. One part is the concatenated sequence of pitch levels for the transcribed tones and the other part is linguistically significant modulations of what Ladd (1992) calls the *tonal space* (see Section 4.3). This partitioning of tunes into tones and tonal space is evident in the Beijinghua utterances in Figure 15.4, where the rising glissando on the word *hé-zi* "box" can be transcribed in terms of a sequence of lower and higher pitch targets ([^{23}] in Chao's tone numbers or LH by Yip's, 1980/1990 analysis) in both cases, but the high target at the end of the glissando in the right-hand utterance is higher because it is realized in an expanded tonal space. This partitioning is also congruent with the distinction in the IPO framework between the parameters that specify the melodically significant glissandi and the parameters that specify a sequence of declination lines for successive phrases.

Carter's account of "the tone-phrasing system of Kongo" is also characteristic in distinguishing between tones that are anchored to specific syllables in a word and tones that signal other phonologically significant anchoring points such as

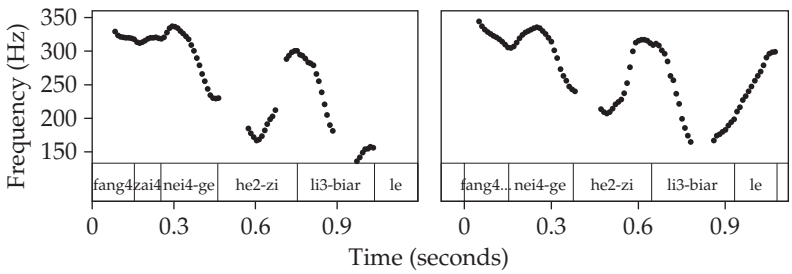


Figure 15.4 Extracted f_0 contours for the sentence *Fàng zài nèi-ge hé-zi lǐ-biar le*. produced by a female speaker of the Beijing dialect of Putonghua in staged dialogues where the context makes it a statement '(I) put (them) in that box.' (left) or a question 'Had (she) put (it) in that box?' (right) (from Lee 2005).

the edges of larger phrases. This distinction is congruent with the distinction between pitch movements on stressed syllables and boundary pitch movements in the IPO models of English and Dutch, as well as with descriptions of some utterance-level melodic contrasts in many languages with lexically contrastive tone shapes. The declarative and interrogative utterances in Figure 15.4 illustrate the Beijinghua case. Whereas the pitch events on earlier syllables reflect the lexically-specified tones, the high tone anchored at the end of the last syllable on the right is a pragmatic morpheme that signals its interrogative speech act.

In this chapter, we will call this lineage of frameworks "prosodic phonology" – using lower case to differentiate this common noun from the homophonous name of two particular frameworks within the lineage (Brazell *et al.* 1966; Nespor and Vogel 1986), and also to make clear that the key ideas invoked by this term are generic. Each of these ideas was developed more or less independently in at least two frameworks within the lineage and was congruent with approaches being developed at the same time in the allied science of speech synthesis. We will adopt transcription conventions from the relevant prosodic-phonology models of tone and intonation systems when we describe f_0 contours of example utterances or discuss melodic contrasts in the language varieties from which these examples are drawn. For instance, the 10 H and L tones in (1) are a possible way to symbolize a subset of the tonal events in the two utterances depicted in Figure 15.1, and the strings of functionally annotated H and L tones in (2) and (3) are possible ways to symbolize the tunes of the three utterances depicted in Figures 15.3 and 15.4.

In both sets of transcriptions in (2) and (3), there are linking lines. Each such line indicates that a tone or tone sequence below is associated to a designated syllable in the orthographic or phonetic transcription above. Also, in both transcriptions, the + infix conjoins tones that are anchored as a glissando around the designated syllable, and in (2), the * suffix on the L+H indicates that English contrasts two rising glissandi, L+H* versus L*+H, which differ in how they are anchored to the designated syllable. By contrast, each % affix in the transcriptions indicates a tone that is anchored at a phrase boundary rather than internally

to a phrase, whereas the suffix in (2) marks a "floating" L phrase tone that is realized somewhere between the preceding L+H* and following H% targets. While we find it useful to adopt these tagging conventions, however, we must emphasize that the symbol strings in (1), (2), and (3) are not narrow phonetic transcriptions. Moreover, they are not even broad phonemic transcriptions until they are construed as names of meaningful configurations of parameter settings in an analysis-by-synthesis model for the speaker's dialect of (1) Japanese, or of (2) British English, or of (3) Mandarin Chinese.

(2) (One day) a royal messenger came to announce a ball.

L+H* L+H* L–H% H* L+H* L–H%

(3) a. Fàng zài nèi-ge hé-zi lǐ-biar le.
[fã.ða ne.ɣ xɤ.ð li.pə.lə]
H+L H+L L+H L L%

b. Fàng zài nèi-ge hé-zi lǐ-biar le?
[fã ne.ɣ xɤ.ðə li.pə.lə]
H+L H+L L+H L H%

4 A Taxonomy of Formal Parameters

In this section, we amplify on the relationships between developments in prosodic phonology and the components of typical synthesis models that allow symbol strings such as the ones in (1), (2), or (3) to be read as pieces of a broad phonemic transcription in some actual or possible formal model of the tone and intonation systems of each of these three languages. We can identify three key developments.

4.1 Segmenting the Melody

The first key development was the notion that the melody of an utterance can be segmented into a string of localized events – single notes or the conjoined notes in glissandi or in more complex sequences such as dipping movements – and that this segmentation is autonomous of the formal properties and functions that allow the native-speaker listener to parse the filter-resonance patterns in terms of the consonant and vowel inventories of a language. This idea gives the name of the autosegmental phonology framework (Goldsmith 1976a, 1979), but it is not unique to that framework. It is implicit in the cataloguing of significant pitch movements in the IPO model, in the identification of turning points in the Lund model (Gårding

1977, 1993), and in the detection of phrase and accent commands in the Fujisaki model (Fujisaki and Sudo 1971; Möbius, Pätzold, and Hess 1993).

It is useful to begin this discussion of the autonomous status of tone segments by reviewing the phonetic bases for the earlier conceptualization of tone as a suprasegmental feature. A great deal of research over the last five decades highlights how a taxonomy of formal properties of vowel and consonant systems across languages emerges from the interplay between information-theoretic principles and the physiology and physics of speech. Indeed, the very fact that there are vowel and consonant systems can be related to the ways in which spoken utterances are naturally segmented by the spectral discontinuities that result when constrictions (i.e. consonant gestures) are superimposed on more open vocal tract postures (vowel gestures). As Goldstein (1989), Ohala (1992), and many others point out, even though the consonant and vowel gestures are not themselves sequenced, the acoustic patterns they produce are effectively sequenced because of two facts about the types of constriction that yield the most robust CV segmentation. First, these constrictions block the transmission of spectral information that gives the listener clues to coarticulated vowel postures behind the place where airflow is impeded. Second, some contrastive features of the most effective consonant segments, such as the place of impedance for a stop, are only audible during the acoustic intervals for coarticulated vowel postures, so that spectral properties at the edges of vowel segments must be treated as transitions between consonant states and vowel states in order to recover these "hidden" consonant features.

Although concurrent tone gestures also are "hidden" by these consonant constrictions, source and filter resonances are to a large extent independently controlled during intervals where airflow is not impeded. As Pierrehumbert (2000) points out, this independence of source and filter for vowels means that tone features are carried on a considerably more separable channel of acoustic information when compared to the "hidden" features of consonants (which, as just noted, are carried on exactly the same channel of spectral resonance patterns that carry the vowel features). Vowel segments are thus the more reliable intervals for transmitting information about concurrent tone gestures during a sequence of consonants and coproduced vowels. This is the psychophysical basis for a type of tone system in which each vowel segment in a word or phrase is the nucleus of a syllable that can be counted off in the metrical structure of the utterance by virtue of its having exactly one associated lexically-specified tone. This kind of "tone syllabification" is especially characteristic of utterances in languages such as Cantonese, where most syllables are monosyllabic content words, and tone features typically preserve the syllable count even at very fast speech rates where the vowel features are swallowed up (see Section 5.1).

Pike (1948: 3–5) reserved the term "tonal language" for a language with this kind of tone-syllabification system, whereas many other later researchers such as Voorhoeve (1973), McCawley (1978), Goldsmith (1984, 1987), and Hyman (2001b, 2006, this volume) defined "tone language" more broadly and used terms such as "restricted tone system" or "pitch accent" to differentiate the "unrestricted tone system" of Cantonese from their accounts of the underlying tone sequence in

languages such as Safwa or Tonga, where words typically are longer and phonotactic constraints strongly restrict what tones can occur where within a word. One of the more fiercely contested questions in prosodic phonology today is the prevalence of tone syllabification as a basis for counting tone targets in the melodic contours of words and phrases. The question arises in part because of the alphabetic bias to model melodic contours of languages with "restricted tone systems" as a succession of "phonetic" tones for all syllables, even over stretches where the observed pitch pattern on the vowels depends completely on the pitch targets for nearby "phonemic" tones.

For example, a common strategy for training models of Mandarin Chinese tone and intonation is to use databases of recorded utterances of sentences such as the two in Figure 15.2 (e.g. Lee, Tseng, and Hsieh 1993; Shih and Kochanski 2000). This strategy in effect treats the language as if it had a Cantonese-like prosodic system. By contrast, Kratochvil's (1998) corpus study suggests that the Beijing dialect at least differs prosodically from Cantonese, and that examples such as the two utterances of the sentence in Figure 15.4 are more typical. The pinyin orthographic transliteration of this sentence in (3) on p. 502 shows nine *zì* (a grammatical unit that Riha 2008 terms the "morpheme-syllable"), but of these nine, *zài*, *-ge*, *-zi*, *-biar*, and *le* are affixes or grammatical particles with "neutral tone" – that is, they have no lexical tone specification, as indicated for the four *zì* other than *zài* by the lack of a tone diacritic. (The locative particle *zài* is transliterated with the diacritic for the [51] lexical tone because it is listed in most dictionaries under the entry for the related locative verb, but the particle and verb are not homophones; the particle has neutral tone and there is no fall in pitch even in the declarative utterance shown in the left panel, where the consonant and vowel are not lenited to nothing, as they are in the interrogative utterance on the right.) How can we account for the f_0 values in neutral tone *zì*?

Chen and Xu (2006) follow recent accounts such as Yip (1980, 1990) to argue that even when there is no lexically-specified tone pattern, each syllable in an utterance has a surface target tone level. The target value for a neutral tone *zì* is M (i.e. midway between the H and L targets of the four contrastive lexically-specified tones) and the actual pitch value is an average of this M with the pitch value of the immediately preceding tone target. By contrast, Li (2003) follows earlier accounts such as Chao (1932, 1968) in assuming that a neutral tone *zì* has no tone target even on the surface. The pitch pattern over such a *zì* can be an extension of the pattern for the last preceding lexically-specified tone (e.g. continuing the fall after the [51] falling tone or realizing the optional rising tail after the [214] dipping tone) or it can be just part of the transition between tone targets on either side (e.g. in (3), the last two syllables are the transition from the L tone on the root morpheme *lǐ* of *lǐ-biar* to the pragmatically-specified L% or H% boundary tone at the end).

The disagreement over which type of account is better is reminiscent of the disagreements that occasionally arise in the literature on transitional elements such as the release phase of obstruents in Berber. Coleman (1998a) transcribes the release as a reduced [ə] vowel, a phone which Dell and Elmedlaoui (1998) and

Ridouane (2008) insist is not part of the phonological inventory of the language, so that by their analyses many syllables are headed by an obstruent consonant, violating a purported universal minimum sonority constraint on syllabicity, albeit not violating most formulations of the universal as a sonority sequencing constraint. However, those disagreements arise very infrequently relative to the consensus view that consonant and vowel gestures in spoken languages tend to be configured syntagmatically in such a way that native speakers and linguists can identify a string of CV units, as in the IPA transcriptions in (3a) and (3b). By contrast, disagreements like the one that gives rise to different counts for the melodic units in these utterances are endemic across the communities of researchers working on tone and intonation. These disagreements are almost inevitable, because they reflect an inherent ambiguity in the parsing of tonal gestures. This ambiguity stems from the fact that a vowel-by-vowel segmentation of the melody is not intrinsic to the production and perception of tone gestures per se, but instead is parasitic on the CV segmentation of the spectral pattern that is intrinsic to the production and perception of obstruent gestures.

Lieberman (1967) proposes a rather different phonetic basis for segmenting the melodic contour that he describes as emerging from the interplay between syntactic structures governing the flow of information in a discourse and the coordination of respiratory and laryngeal postures to control expiratory airflow for phonation. In particular, he suggests that, absent a "marked" gesture to change laryngeal tension, the posture for sustained phonation results in a rise to high f_0 at the beginning of controlled expiration and a rapid fall in f_0 toward the end, as in the combination of prominence-lending movements in the IPO "hat pattern" in Figure 15.3. This rise and subsequent fall forms a natural unit of segmentation, which Lieberman calls the "unmarked" breath group. He also describes a "marked" breath group, which instead has a final rise produced by a localized laryngeal-tensing gesture. He claims that a comparably localized gesture to boost subglottal pressure can also make for a more extreme early rise or a rise in other positions to mark focal prominence ("emphasis" or "contrastive stress") in the discourse context of the utterance.

Although Ohala (1970), Collier (1975), and others discredited Lieberman's characterization of a definitive role for subglottal pressure in the production of local melodic events such as the L+H* rise when used to mark focal prominence on a particular syllable or word in English, Lieberman's depiction of an early rise and late fall that defines the melodic contour for an "unmarked breath group" captures a fairly common aspect of phrasal melody. Safwa, Basque, Japanese, French, and many other languages use a small set of tone sequences, often involving a rise in pitch anchored near the beginning and a fall in pitch anchored later, to highlight the edges of utterances and to segment them into smaller prosodic phrases. These same prosodic phrases often seem to be the domain for specifying an expanded or compressed tonal space to express the relative prominence of the constituent as a whole (see Section 4.3), and it is less implausible that this expression of phrasal prominence relationships could involve adjustments to the pulmonary expiration rate as a mechanism for overall volume control.

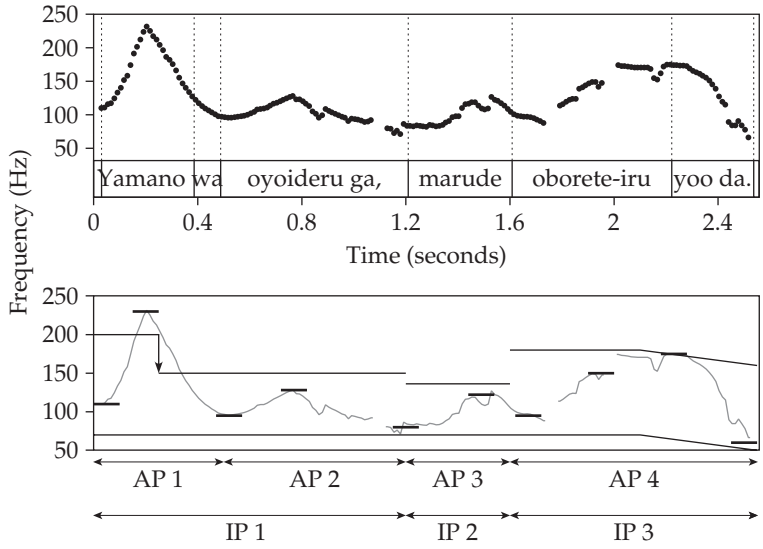


Figure 15.5 Extracted f_0 contour for an utterance of the sentence in (3) meaning 'YAMANO is swimming, but he's nearly drowning right now.' produced by a male native speaker of Tokyo Japanese. The copy in the bottom panel shows overlaid lines for tone targets and tonal space settings that are described in Section 4.3. (See Figure 15.1 for a spectrogram of this utterance.)

The utterance in Figure 15.5 (which repeats the f_0 contour from utterance (b) in Figure 15.1) illustrates this delimitative aspect of the tone-phrasing system of Japanese. There are four prosodic phrases, each of which is marked by an initial rise in pitch. This rise is analyzed in the X-JToBI tagging conventions as a sequence of a low boundary tone that is anchored strictly at the phrase edge and a high phrasal tone which is timed to follow the low tone at some loosely fixed distance which depends both on the prosodic structure of the first syllable and on distance to the next melodic event. In every phrase but the third, there is also a steep fall at a designated syllable that is marked by an apostrophe in the transliteration of the word in (4a). This lexical specification for anchoring a HL tone sequence at some designated syllable differentiates "accented" words such as *oyo'ideru* 'is swimming' from "unaccented" words such as *oborete-iru* 'is drowning'.[3] Although the verb form in the second phrase is unaccented, there is a steep fall in this VP because the following evidential particle is lexically accented. Even in phrases that contain no accented words, however, there is most typically a fall, albeit often a more gradual one, which the X-JToBI conventions analyze as a transition from the high target at the end of the phrase-initial rise to the low boundary tone at the following phrase edge. Prosodic groups (*accentual phrases*) can be counted off in an utterance from the distribution of the phrasal rises and subsequent steep or gradual falls. Also, while the tone patterns differ in other dialects, with some

having more complicated lexical contrasts and some having no lexical contrasts, the metrical structures that are defined by the distribution of tones relative to accentual phrase (AP) boundaries seems to be shared across dialects.

(4) a. Ya'mano-wa oyo'ideru-ga, marude oborete-iru yo'o-da.
[ja.ma.no.ɯa o.joi.de.ɾɯ.ŋa maɾ.ɯ.de o.bo.ɾe.te.i.ɾɯ jo:.da]
%L HL L% HL L% H– L% H– HL L%

b. Ya*mano-wa oyo*ideru-ga, marude oborete-iru yo*o-da.
H*L L H*L L H L H*L

c. [ja.ma.no.ɯa o.joi.de.ɾɯ.ŋa ma.ɾɯ.de o.bo.ɾe.te.ɾɯ jo:.da]
[H L L L L HL L L L L H H H L H H H H HL L]

Accounts such as Kawakami (1957) and Pierrehumbert and Beckman (1988) focus on the way that the tone patterns mark off the salient prosodic groups to make for the pan-Japanese metrical system. In such accounts, melodic contours for utterances in the standard dialect are segmented only into those tones that are anchored relative to the phrase edges and those tones that are anchored at the designated syllables of accented words. All other parts of a contour are described as tonally "underspecified" and modeled as transitions between the nearest tones on either side, making tonal transcriptions of Tokyo Japanese utterances such as the one shown in Figure 15.5 look like transcriptions of utterances in the Autosegmental-Metrical model of the American English intonation system that was invoked in (2). The transcription in (4a) illustrates this, using the X-JToBI conventions (Maekawa *et al.* 2002).

By contrast, in Kindaichi (1957) and Haraguchi (1977), the focus is primarily on making a spare underlying representation for the lexical contrasts between the absence versus presence of the HL sequence and (if present) among different anchoring positions within the word. These contrasts are represented by marking the designated vowel with a * to show where the HL sequence is to be inserted at the initial stage of deriving the surface pitch pattern. The pattern on other parts of the accentual phrase is modeled by derivational rules that conditionally insert L and H tones on the initial and final vowels, as in (4b), and then copy the inserted tones or the lexically-specified tones onto other vowels, to produce a "fully-specified" surface pattern, as illustrated in (4c).[4] This tone-spreading account makes the intonation system of Tokyo Japanese look like Voorhoeve's (1973) picture of the "restricted tone system" of Safwa and also like Goldsmith's (1984) account of "tone and accent in Tonga" a decade later.

The difference between the 13 tone segments assumed in the transcription in (4a) and the 21 tone segments assumed in the transcription in (4c) is also parallel to the difference between specifying tones for just four of the *zì* in Li's (2003) model that yields the transcription in (3a) as compared to specifying these four

plus the five M targets for the neutral tone *zì* in Chen and Xu's (2006) model. In both of these cases, one account assumes that the sequence of syllables (or other potential tone-bearing units) is "fully specified" for tone targets whereas the other account assumes that the nodes at this level of the prosodic hierarchy are "underspecified" for tone. These names characterize the disagreement in terms of their different assumptions about the set of localized pitch events – that is the first of the three synthesis model components listed in Section 3.3.

Such disagreements have consequences for the depiction of the "underlying" tone specification. For example, in the fully-specified account of Japanese tone patterns, the starred tone of the H*L word melody is associated to the designated mora (which is marked with a * in the lexicon) at the first stage of the derivation, and then a L is inserted on the first tone-bearing unit of an AP just in case that mora does not already bear a tone specification. This account therefore predicts that there will be no tone difference between a sequence of clauses such as *yonde-mi'ru* 'call and then see' and a verb-auxiliary construction such as *yonde-mi'ru* 'try calling' since in both cases the initial vowel of *mi'ru* will already have an associated H tone at the stage of the derivation when an initial L is conditionally inserted, as in (5c). By contrast, in the underspecified account, that first L is a boundary tone that marks the edge of an AP whether or not the first syllable is accented. Thus, utterances of the two-clause sequence often would be distinct from utterances of the verb-auxiliary construction, because the two-clause sequence often will be produced as two AP, as shown in (5a).

(5) a. yonde mi'ru 'call and then see' yonde-mi'ru 'try calling'

%L H– L% H+L L% %L H– H+L L%

b. [jon.de mi.ɾɯ] [jon.de mi.ɾɯ]

c. yonde mi*ru 'call and then see' yonde-mi*ru 'try calling'

L H H*L L H*L

d. [LH H H L] [L H H H L]

It is important to note that these differences in the analysis of the underlying forms stem from the more fundamental disagreement about the nature of surface phonetic representations. In the X-JToBI account of Tokyo Japanese (as in all ToBI framework accounts), the surface phonetic representation is the actual pitch pattern, as deduced from representations such as Figure 15.5, which shows an f_0 track calculated from a recording of a specific utterance of (4a), or as shown in (5b), which is a schematic "pitch track" summary of the many f_0 contours that we have observed for actual utterances of the phrases in (5a). In Haraguchi's account, by contrast, the surface phonetic representation is still a symbolic transcription – a

sequence of discrete pitch targets associated vowel-by-vowel, as in (5d). On the surface, then, this account makes Japanese look like an unrestricted tone language. How can we decide between these two accounts?

Pierrehumbert and Beckman (1988) made the following predictions. If the fully-specified account is an accurate representation of what the speaker intends to produce, then a sequence of spread L tone targets (as in the last four vowels in *oyo'ideru-ga* in (4c)) or a sequence of spread H tone targets (as in the second through sixth vowels in *oborete-iru-yo'o-da* in (4c)) should show the same pattern of actual f_0 values over the associated vowels regardless of the length of the sequence. In the underspecified account, by contrast, the f_0 contour over such stretches could fall or rise at different rates, depending on the distance between the two tone targets specified at the surface. Pierrehumbert and Beckman tested these predictions using a set of elicited utterances of three-phrase sentences in which both the accent status and the number of syllables in the words in the middle phrase were systematically varied. For unaccented medial phrases, they measured the slope of the f_0 downtrend over the interval between the peak f_0 near the beginning of the accentual phrase to the minimum f_0 at the next phrase boundary – that is, over an interval that would be represented as a sequence of H tones in Haraguchi's account but as a mere transition from a phrasal H– to a L% boundary tone in the underspecified account. For accented phrases, they fit two slopes, differentiating the steep fall of the H+L tones at the designated syllable (which they predicted to have a fixed duration and slope) from the shallower decline over the variable-length region from the L of the accent to the L% at the end of the phrase. In both cases, the slope of the downtrend over the variable-length region up to the phrase edge was steeper for shorter intervals and shallower for longer ones, as predicted by the underspecified account.

In differentiating between the fully-specified and underspecified accounts of Tokyo Japanese tone patterns, Pierrehumbert and Beckman (1988) fit very simple (straight-line) curves to the f_0 contour over both types of tonally unspecified intervals. As Pierrehumbert (1980: 12), van den Berg, Gussenhoven and Rietveld (1992), Beckman and Pierrehumbert (1992), Myers (1998), Ladd and Schepman (2003), and many others point out, however, the shape of a transition over tonally unspecified regions is a research question in its own right. Moreover, it is a question that is tied up inextricably with questions about alignment or anchoring – that is about how the speaker synchronizes tone gestures with vowel and consonant gestures so that the listener correctly parses where the targets are anchored in relation to prosodic positions such as stressed syllables and phrase boundaries.

4.2 Anchoring the Tones in Time

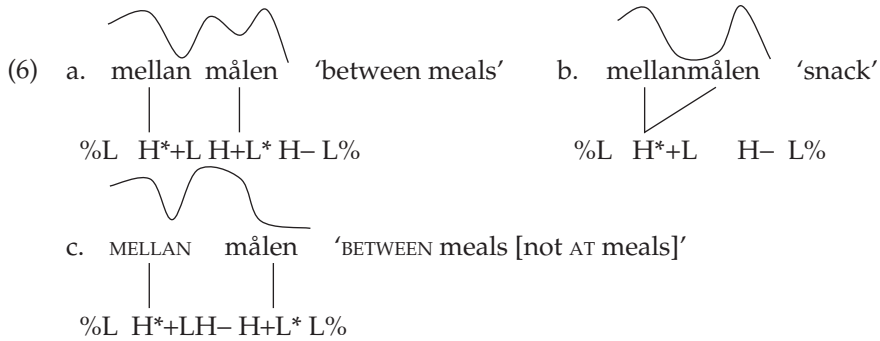
The second key development in prosodic phonology was the idea that tonal autosegments are not suprasegmental features of the vowel segments on which they realized. Rather, vowel (and consonant) segments as well as tone segments are associated to positions in a metrical structure, and this structure and the association patterns are objects of study in their own right. This idea is often associated

with what is termed the metrical phonology framework (e.g. Liberman 1975, 1979; Liberman and Prince 1977; Selkirk 1981b), but again, it is not unique to that framework. It is developed more fully in the treatment of coarticulation of consonant and vowel features in what is called the articulatory-phonology framework (see Browman and Goldstein 1986; Byrd 1993; Byrd and Saltzman 2003, other work reviewed by Fletcher 2010). For tonal autosegments, this idea is implicit in the functional separation between prominence-lending pitch movements and boundary pitch movements in even the earliest IPO system models, and it corresponds to the distinction between turning points and pivots in the Lund model and to the distinction between phrase commands and accent commands in the Fujisaki model.

To show how this development was separate from the first key idea, we begin by comparing what "association" means in the two different accounts of Japanese discussed above. The fully-specified "phonetic representations" in (4c) and (5d) can do away with the link lines and just list the string of H and L tones, reflecting the assumption that each tone is aligned simply to coincide with the vowel or moraic nasal segment to which it associates by rule. Beckman, Hertz, and Fujimura (1983) describe a synthesis model couched in this fully-specified autosegmental phonology framework, which specifies a target f_0 value for the H or L tone midway through each vowel or moraic nasal in this way. By contrast, the underspecified surface transcriptions in (4a) and (5a) must show link lines to identify the accent tones and their designated syllables in accented words. Other tones must be annotated for their anchoring relationships. The annotation conventions differentiate the %L and L% boundary tones that anchor tightly at the phrase edge from the H– phrase tone that is only loosely aligned relative to the edge of the accentual phrase that begins with an unaccented word. Pierrehumbert and Beckman (1988) describe a synthesis program couched in the AM framework which specifies target f_0 values at various time points that are chosen to relate the linguistically significant f_0 peaks, valleys, and inflection points ("elbows") in the phonetic representation in Figure 15.5 to the functional differences among the accent tones, the boundary tones, and the phrase tones. Although the input is a sequence of tones, the ways in which tone sequences such as the L% H– are anchored to positions such as the phrase edge makes their model much more like Kawakami's account than like Kindaichi's.

Pierrehumbert and Beckman's (1988) model of Japanese tone structure relied crucially on Bruce's seminal model of Stockholm Swedish tone patterns (Bruce 1977, 1982, 1987, 1990). In Bruce's model, there are three types of tone which are anchored in different ways to designated constituents or positions at several levels of a hierarchy of prosodic units. The first two relevant levels are the grouping of consonant and vowel constituents into short (unstressed) and long (stressed) syllable constituents, and the grouping of unstressed syllables together with neighboring stressed syllables into word constituents. The second level is marked tonally by the *word accent*, a H+L tone sequence that is anchored to a designated strong syllable in each word. This culminative distribution of the H+L sequence means that in longer Swedish utterances, words can be counted off in the melodic contour for an utterance by recognizing the word-accent tones and their anchoring points. Above the word level, whole utterances and prosodic phrases within

utterances are delimited by boundary tones such as the L% for the "terminal juncture fall" (Bruce 1983: 223). Also, the melodic contour for each phrase must include a H– tone, called the *sentence accent* in Bruce (1977), the *phrase accent* in Pierrehumbert (1980), and the *focal tone* in Gussenhoven and Bruce (1999). The focal tone is realized just after the word accent of the word with "sentence stress" – that is a word that is in narrow focus in the discourse context or the last word in the phrase when there is broad focus over the whole phrase. All the tone types are shown in the sample transcriptions in (6). These schematic "pitch contours" are based on f_0 tracks given in Bruce (1977) and are intended to give a sense of the typical patterns of truncation and undershoot.



As in most dialects of Swedish, the Stockholm variety has a lexical contrast between two anchoring patterns for the word accent, transcribed by Bruce (1990) as H+L* ("Accent 1") versus H*+L ("Accent 2"). In Accent 1 forms such as *anden* 'the duck', *anamma* 'accept', and *målen* 'meals' produced in contexts with one or more preceding syllables, the H+L* denotes a fall to a low pitch target within the stressed syllable that starts from a pitch peak or a high inflection point (an "elbow") about 120 ms before the low target. In Accent 2 forms such as *anden* 'the ghost', *lämna* 'leave', and *mellan* 'between' there is a peak or high elbow within the stressed syllable and a fall to a valley or a low elbow 120 ms later. A compound word such as *mellanmålen* 'snack' is marked by a H*+L (Accent 2) anchored to the designated syllable of the first component and no word accent on any later component. This accenting in compound words mirrors the typically initial stress in the native Germanic stratum of the lexicon.

Other important concepts are truncation and undershoot. When an Accent 1 word with initial stress is initial in its utterance, the leading H of the accent will be effectively "hidden" by the preceding silence, so that the underlying H+L* is truncated to be just the L* target on the designated syllable. Also, when an Accent 1 word with final stress is final in its phrase, the close succession of fall for the H+L* followed by rise to H– and fall to L% leaves very little room for the word accent to be realized. There is undershoot so that the L* is effectively a mid tone. By contrast, the trailing L of the Accent 2 fall is typically fully realized, because the designated syllable in an Accent 2 word cannot be final. Moreover, the duration of the transition from the H* target to the elbow for the trailing L is very stable. At the other extreme, the H focal accent has no very fixed constraints

on its alignment other than that it is after the accent tones of the focalized word. In compound words, it is especially late, because the trailing L of the word accent has a secondary anchoring point at the stressed syllable of the second (or last) word in the compound. This account of the focal H– as a "floating tone" is invoked in AM-framework transcriptions by showing no line linking it to a designated syllable.

Two aspects of Bruce's work are especially noteworthy. The first was his rigorously controlled laboratory phonology methods. He designed his materials to allow a systematic comparison of melodic contours for Accent 1 and Accent 2 words of different lengths in both final and non-final position and in both non-focal and narrow focus contexts. This was necessary for him to be able to disentangle the tones that are specified by the lexicon from the tones that mark other levels of prosodic organization. He used analysis by synthesis to verify the segmentation of the melodic contour into these disparate elements and to examine their timing relative to the consonants and vowels at phrase boundaries and at the designated syllables within each phrase.

An equally important aspect of his work was his rigorously imaginative adaptation of key ideas from prosodic phonology. He did not let broad-stroke typologies dictate what analogies could be drawn between the tone patterns of Swedish and the intonational accents of English, and was among the first to grasp the implications for prosodic phonology of Bolinger's (1958) theory of pitch accent in English as interpreted by Vanderslice and Ladefoged (1972). He saw that the syllable bearing the word accent is not the only potential site for anchoring a tone target in a citation-form utterance of a Swedish word, and that tones realized at variable distances from the accented syllable in many dialects (including the Stockholm one) might reflect rhythmic organization above the word. This let him re-conceptualize the originally simpler theory of "association" only between autonomously segmented tones and some unspecified temporal location within the set of "tone-bearing-units" at just one relatively low level of the prosodic tree (that is, either to just the V nodes, just the μ nodes, or just the σ nodes) as a more complex synchronization at "critical timing points" (Bruce 1983: 234) that speakers and listeners control to resolve potentially conflicting demands in different phonological domains.

For the speaker, these conflicts involve "the interaction between the timing of phonatory and articulatory gestures" (Bruce 1983: 222), which cannot follow an invariant rhythm because the words in a sentence can be one syllable or longer, initially-accented or accented on a later syllable, in focus or subordinated to a neighboring constituent, and so on. Consonant and vowel gestures in a particular utterance of a string of words must be synchronized with each other so that the listener can parse the syllable count, hear whether each syllable is stressed or unstressed, and, if stressed, whether it is an extended vowel gesture or a coda consonant gesture that contributes the second mora in the syllable. Tone gestures also must be synchronized with the consonant and vowel gestures so that the listener can hear which stressed syllables are accented, whether an accent is H+L* or H*+L, and what word is highlighted by the focal H–. These different prosodic

functions impose different demands. Realizing the word-level contrasts between short and long syllables and between H+L* and H*+L accents places stringent demands on the timing of the targets. In realizing the utterance-level contrast between focus and background, on the other hand, the exact timing of the focal H– is less relevant than achieving a particular target peak value, since the latter signals prominence relationships among words and phrases as well as among syllables within each word. Conflicts among these demands can be reconciled by adapting the tone targets (e.g. through truncation or undershoot) or by adapting the vowel and consonant targets (e.g. lengthening a final accented syllable to realize a complex sequence of word-accent tones, focal H–, and boundary tones, as suggested by Lyberg 1981). To model the relevant interactions, the segments and tones must be observed in more contexts than citation-form utterances.

In Bruce's original formulation of this "synchronization hypothesis" he differentiated between two orientations for evaluating the synchronization. From the "phonological point of view" of a "production-oriented model" it is useful to specify the critical timing points for the underlying tone targets, but these may not map neatly onto the "perceptually critical" f_0 events such as rising or falling glissandi. For example, in his own perception experiments on the timing of the H+L* targets of Accent 1 versus the H*+L targets of Accent 2, Bruce found that the times of the starting and ending points traded off with the steepness of the fall in a way that suggested that subjects listened for the point of maximum velocity in the middle of the fall. However, reference to this midpoint time "is possible only in a sonorant environment" so that "from a perceptual point of view, it is probably an estimate of the timing of the entire f_0 change . . . that is decisive" (Bruce 1983: 231).

Bruce's hypothesis was developed to account for the variable realizations of Swedish word accents across different sentence contexts and different dialects, but it was an important precursor to the AM model of Japanese tone structure presented in Pierrehumbert and Beckman (1988), as well as to the development of the AM framework generally. Initially, development of the framework was addressed more to the production-oriented aims of finding "invariant" or "underlying" tone targets and their modes of association to phonologically defined positions in the hierarchy. For example, Pierrehumbert and Beckman (1988) proposed that the L% and H– tones in their model of Japanese are associated initially to the two accentual phrases on either side of the boundary that the pitch rise marks, but then that each tone is also associated secondarily at a later derivational stage to the first unaffiliated mora in the accentual phrase that begins at the boundary. They observed differences in the shape of the rise and in measured f_0 minima for what they called the "strong L%" versus the "weak L%" and they attributed these differences to a contrast between having and not having a secondary association to the first mora in the following accentual phrase. Gussenhoven and Bruce (1999) similarly propose to account for the shape of the trough in citation-form utterances of compound words in Stockholm Swedish in terms of a secondary association of the trailing L of the H*+L accent. Grice, Ladd, and Arvaniti (2000) catalogue other examples of languages where a phrase accent can

be analyzed as having a dual affiliation to both the edge of a larger prosodic domain and to some designated syllable within the domain. This focus on tone targets and their anchoring relative to the prosodic structures that the speaker controls is congruent with the production-oriented approach of the articulatory-phonology framework (e.g. Browman and Goldstein 1990; Byrd 1993). For example, Xu and Liu (2007) apply a model of Putonghua lexical tone alignment to examine syllables of both Putonghua and English in order to probe for universal patterns in how an onset consonant gesture is anchored to its syllable to be coarticulated with the relevant vowel. This application suggests some of the questions that can be addressed fruitfully using production-oriented models that assume invariant underlying tone, vowel, and consonant targets for the speaker that are aligned with each other to reflect the "temporal signature of prosody" (see Fletcher 2010).

Other recent work, however, suggests that the time is ripe to begin to reorient our models to incorporate constraints on the listener, too. For example, Arvaniti, Ladd, and Mennen (1998, 2000) show that the timing of prenuclear rising accents in Greek does not fall out from a simplistic model that designates either the L or the H as the target that is associated to the designated syllable. Rather, the L is anchored just before the syllable-initial consonant and the H is anchored to coincide with the CV boundary in the following syllable. Unlike Swedish, Greek has only five vowels, with no prosodic contrast between short and long vowels or short and long consonants. Many syllables are CV and vowels tend to be quite short. Also, whereas many Swedish words follow the common Germanic pattern of root-initial stress, the position of the stressed syllable in a Greek word is constrained to occur only on one of the last three syllables. Within this three-syllable window, stress placement is "phonologically unpredictable" (Arvaniti 1999: 171). Given these characteristics of the language, the observed anchoring pattern for Greek prenuclear rising accents may have emerged as a way to provide the listener with a robust "estimate of the timing of the entire f_0 change" in order to reliably parse the location of each accented syllable in an utterance.

These demands on the Greek listener are different from the demands on the listener to a language such as Dutch, where there are many more than five vowels in the inventory; vowels are typically longer, and there also is a much larger variety of typical syllable structures, including a contrast between syllables with short vowels and syllables with long vowels. Ladd, Mennen, and Schepman (2000) show that in Dutch, the timing of the end of a rising accent is not fixed in the same way as in Greek. Rather, it is later relative to the end of a syllable with a short vowel and earlier relative to the end of a syllable with a long vowel, and this difference in anchoring of the endpoints supports the vowel length contrast even for speech rates and discourse contexts where the vowel durations themselves are not robustly different.

Arvaniti, Ladd, and Mennen (2000) end their paper with a call for more research both to refine what "association" means for our models of the prosodic structures that the speaker intends to produce and to devise better methods for understanding how targets and their timing properties are realized in the speech signal that the listener parses. One promising line of research in this vein is comparative

work such as Smiljanić (2006). Smiljanić looked at accent-related rises in standard Serbian and Croatian, language varieties which are mutually intelligible but which differ in whether there is a lexical contrast between word accents with an early versus a late peak. She found that speakers of both varieties signal focal prominence on a word by manipulating the timing as well as the maximum f_0 value of the pitch rise to the accent peak. However, the timing effect is much smaller in Serbian, where the anchoring of the rise also signals the contrast between the two word accent types. We need more such comparative work in other language groups to develop our understanding of the potential role of functional load in the interaction between demands on production and demands on perception. We also need more work that does what both Bruce (1977) and Smiljanić (2006) did – namely, to observe tones in words across a good variety of sentential and discourse contexts, to see how variation in the demand for precise "horizontal" anchoring of tone targets relative to critical positions within a word interacts with variation in the demand for precise "vertical" positioning of the tone targets relative to the tonal space.

4.3 Tone Scaling and the Tonal Space

The third key development in prosodic phonology was the idea that speakers can raise or lower and expand or compress the local tonal space as a whole and also independently scale tone targets up and down within the tonal space, to reflect both autosegmental contrast and relative metrical strength, as well as other sorts of linguistic (or "paralinguistic") relationships. While there is a broad consensus that this separation of "vertical" position into two parts is necessary, the separation is realized differently in different AM-framework models, and the linguistic nature and formal status of the independence remain controversial. We illustrate one way in which the independence of tone scaling and tonal-space specification parameters has been modeled, using an AM-framework description of Tokyo Japanese that was first developed and tested in an analysis-by-synthesis system by Pierrehumbert and Beckman (1988) and then modified by Maekawa *et al.* (2002) in developing the X-JToBI conventions that were used in tagging the Corpus of Spontaneous Japanese (Maekawa 2003). The separation of parameters in the model corresponds roughly to the specification of variable accentuation levels for turning points independent of the parameters of the tonal grid in the Lund model of various dialects of Swedish and to the independent specification of amplitude values for accent commands and phrase commands in the Fujisaki model of various dialects of Japanese.

As noted earlier, we have adopted Ladd's (1992) term "tonal space" to talk about the effects that the IPO-framework models generate by specifying variable starting values and slopes for declination lines over different parts of an utterance. Ladd chose this term to have a framework-neutral way of referring to what Chao (1930) called the pitch "range" when he proposed his "system of tone letters" and the corresponding numerical notation that we used to indicate the lexically contrastive pitch pattern on each of the syllables in the transcriptions of the

Putonghua utterances in the caption to Figure 15.2. Chao (1932: 124) identifies "several abstractions" that must be made to record the pitch patterns that differentiate the tone classes in any dialect of Chinese. Specifically, each pitch level must be calculated "relative to the speaker's range of voice, so that what would be a low tone for a soprano is actually higher in pitch than the high tone of a tenor." Moreover, "the range of pitch between different tones and within the limits of moving tones is also a variable quantity depending on force of articulation and force of vocalization."

The abstraction over different speakers' voices is analogous to the abstraction over different vocal tracts when computing targets in some speaker-normalized representation of the vowel formant space. The abstraction over variable "force of articulation" is analogous to the constancy of vowel-class identity across the hyperarticulation-hypoarticulation continuum (Lindblom 1990). An important difference between these two spaces is that the "force of articulation" and the "force of vocalization" effects on vowel formant values are necessarily small compared to speaker effects, because maneuvers such as contracting the strap muscles to lower the larynx can change a soprano's vocal tract length by only a small amount relative to the typical length difference between her vocal tract and a tenor's. By contrast, the "force" effects on pitch values can be extremely large relative to the differences across speakers, so that the soprano's H tone in a very subdued speaking style can be much lower than the tenor's H tone in a very forceful speaking style. A phonological consequence of this difference between the phonetic spaces is that when force of articulation and force of vocalization effects on vowel formant values are phonologized as linguistically significant markers of strong versus weak positions in the prosodic hierarchy of a language, the markers typically can be described in terms of a small number of discrete prosodic constraints on what vowel targets can be specified for moras or syllables in different positions of the hierarchy. Analogous prosodic constraints on what tone targets can be associated in different positions are fairly common across spoken languages (cf. Section 5.3), but an even stronger universal is the phonologization of the control parameters for positioning tones within the tonal space and for varying the dimensions of the tonal space itself so that these can act not just as discrete markers of the set of categorical contrasts in prosodic organization, but also as gradient markers of more subtle differences in relative metrical strength as well as of other linguistic scales.

The bottom panel of Figure 15.5 illustrates the parameterization of the tonal space and of tone scaling that Pierrehumbert and Beckman (1988) built into their synthesis model for Japanese, as these parameters are understood in the version of this model that was incorporated into the X-JToBI labeling conventions on the basis of later research that is reviewed in Venditti, Maekawa, and Beckman (2008). In this model, there are tonal-space or tone-scaling effects that refer to three different types of prosodic constituent – the intonation phrase (IP), the accentual phrase (AP), and the prosodic word (ω).[5]

At the beginning of the first IP, the reference line that defines the bottom of the tonal space is initialized to reflect overall engagement or volume within the

speaker's voice range. The reference line for the utterance in Figure 15.5, for example, is initialized at 70 Hz. This value is maintained until late in the last IP of the utterance, where the effect of "final lowering" kicks in, to signal discourse-level functions such as topic shifts or yielding of the floor to the other speaker. Final lowering is a change in the reference line time function, from having a fixed value to showing a decline over some span at the end of an IP. In the turn-final utterance shown here, for example, the effect reaches in to lower the reference line by 44 Hz per second starting at 0.45 seconds from the end of the last IP.

The IP is also the level of prosodic structure where the value for the top of the tonal space is (re)initialized. The initial topline values for the three IPs in the utterance in Figure 15.5, for example, are set at 130, 66, and 110 Hz above the reference line.

The IP is also the domain of downstop, a compression of the tonal space triggered at each lexical accent. This effect is implemented in the model by reducing the distance of the topline from the reference line by a fixed ratio. The downstep ratio in the first IP that is triggered by the accent on the first syllable of the prosodic word *Yamano*, for example, is 0.62 – compressing the tonal space to 62% of its original span.

Tone targets at the level of the IP, the AP, and the ω are then positioned within the local tonal space that is defined by the additive effects of the initial IP topline specification, the compression at each previous downstep, and edge effects such as final lowering. Position within the tonal space first of all defines the discrete contrast between H tones (the targets that are closer to the topline) and L tones (the targets that are closer to the reference line). The level of the AP, for example, is defined by the rise from the %L or L% boundary tone to the H– phrase tone. At the level of the ω, the lexical contrasts among accented and unaccented words are expressed by the presence and (if present) the location of the H+L accentual fall.

The top and bottom of the tonal space also act as a reference for continuous within-category contrasts in metrical strength. Stronger L tones are scaled lower, to be closer to the reference line, and stronger H tones are scaled higher, to be at (or even above) the topline. Some of these strength contrasts are intrinsic to the tone target type. Within an AP containing an accented ω, for example, the H tone of the H+L word accent is intrinsically stronger than the H– of the phrase-initial rise; it will be higher relative to the topline, other things being equal. Other strength contrasts are extrinsic and reflect other types of linguistic structure, such as the discourse-level differentiation between given and new information. The imagined context for the performance of the utterance in Figure 15.5, for example, is a conversation between two spectators at a triathlon relay race. The other speaker has just asked whether the athlete who is swimming could be Yamano. The H of the word accent in the first AP goes above the local topline to reflect the discourse-level prominence of *Yamano* as a contrastive topic. The H of the word accent in the second AP is much lower, reflecting both the compression of the tonal space at the downstep and also the given status of the verb *oyoideru* 'is swimming' in this dialogue context.

The effects that are illustrated in Figure 15.5 are parameterized in somewhat different ways in the Fujisaki model that Hirai, Higuchi, and Sagisaka (1997) used

to analyze several large multi-speaker corpora in order to develop the intonation component of CHATR, a concatenative speech synthesis system with prosody-based unit selection (Campbell and Black 1997). For example, in the Fujisaki model, downstep is not modeled explicitly, but instead falls out from the choice of amplitude values for successive accent commands. At the same time, there are important commonalities between these two models. In particular, both models encode relative prominence relationships among tone targets using two different sets of parameters. In the Fujisaki model, there is a step function (the phrase command) to (re)initialize the backdrop tonal space at the beginning of each new IP and a matched pair of step functions (the accent command) that generates the rise to the first H target in each AP (as well as the fall at the accent or at the end of the AP if there is no accented W in the phrase), and the amplitude of each of these two commands is a continuously variable parameter. That is, the distinction between these two amplitudes corresponds roughly to the distinction between the tonal space parameters that are initialized at the level of the IP and tone scaling parameters that are specified for the tone targets that are obligatory at the level of the AP in the AM-framework model depicted in Figure 15.5.

One critically important difference between the two models is the treatment of L-tone scaling. As noted already, prominent L tones are scaled downward toward the reference line in this AM-framework model. In the utterance in Figure 15.5, for example, the L% tone at the IP boundary after *oyoideru-ga* is lower in the local tone space than the L% tone at the mere AP boundary after *Yamano-wa*, reflecting the difference in metrical strength between those two positions in the prosodic hierarchy. In Osaka Japanese, where there is a contrast between %L-beginning and %H-beginning words, there is a similar downward scaling of this initial L tone as well as a delay in the beginning of the following rise in L-beginning unaccented words under focal prominence (Kori 1987) and in pragmatically loaded questions (Miura and Hara 1995). These effects would be difficult to model in the Fujisaki framework without introducing another type of (downward pulsing) accent command that can be positioned at places other than the beginning of the AP.

Another critical difference is in the treatment of effects such as final lowering. Beckman and Pierrehumbert (1986) suggest that extreme final lowering defines one end of a continuum which has (at the other end) an effect that they call "final raising" which they observed in syntactically unmarked questions. As already stated, in the AM model in Figure 15.5, this kind of edge-in effect is modeled directly as a change in the shape of the tonal space at the end of some phrasal grouping, analogous to the way that phrase-final lengthening and initial strengthening are treated in the π-gesture model of Byrd and Saltzman (2003) and other articulatory-phonology framework models (see review in Fletcher 2010). In the Fujisaki framework, by contrast, such edge-in upward or downward slope differences cannot be modeled directly. There is a necessary downtrend across the tonal space for the whole IP, because the phrase command impulse is smoothed by a filter function and the resulting curve is convolved with the concurrent accent commands, each of which is also smoothed by a different filter function. However, since these filter shapes are intended to reflect "hard" physiological constraints

(cf. Öhman 1967; Fujisaki 1983), they are not under the speaker's direct control. In order to vary the slope as a way of marking structural properties such as the discourse property of being turn final, the modeler must insert a phrase command with just the right amplitude at some place near the end to counter the downtrend from the damping function. The inability to model systematic slope variation directly makes the Fujisaki model fundamentally different not just from the AM-framework model of Japanese, but also from Grønnum's very different model of functionally similar effects in Standard Danish (Thorsen 1983, 1985, 1986).

Grønnum's model, on the other hand, is fundamentally different from both the Fujisaki model and the AM model in that all aspects of the tone pattern are treated in terms of a hierarchy of trend lines, with slopes that are specified for the nested spans of the individual stress groups within individual clauses in a semantically coherent text. Factors affecting these slopes are the length of the span (e.g. the clause-level slope is very steep for clauses containing fewer stress groups and very shallow for clauses containing many stress groups) as well as the same discourse-level factors that Beckman and Pierrehumbert (1986) identify as the function of edge-in effects such as final lowering.

As Grønnum (1990: 199) points out, the Danish effects are formally distinct from the Japanese effects in that they are global and not localized to the phrase end. The downtrend that signals finality "does not just reach one half-second in from the end, it reaches in all the way back, across several . . . stress groups to the onset of the utterance." In Liberman and Pierrehumbert's (1984) AM-framework re-interpretations of Grønnum's results, this longer-range clause-level slope function is modeled in terms of downstep triggered locally at each successive accent. The speaker would have to be able to specify a different downstep ratio at the beginning of each IP in order to simulate the difference in slope between a final and non-final clause. The even longer-range slope of the text, on the other hand, is modeled in terms of the speaker's specific choices for reference line or topline initialization values for the successive phrases. Grønnum (see Thorsen 1984: 307) criticizes this "local" approach as arbitrarily allocating responsibility to disparate sets of formal parameters to account for the functionally uniform hierarchy of syntactic and semantic coherence.

Grønnum (1995: 348) voices a related criticism in her review of the equally "local" treatment of downtrends in Möbius's (1993) Fujisaki-framework model of German. Specifically, she points to his results that the amplitude of the phrase accent command depends on the number of accents and also (in short sentences) on whether the accent is early or late. After quoting from Möbius's comparison between his more "local" approach and her "hierarchical" one, she agrees:

> That is exactly . . . the problem with FUJISAKI's model as adapted by the author: it permits phrase command amplitudes to depend on accent location, and it does not supply criteria for a principled choice between several sets of phrase and accent parameters which each render an acceptable f_0 copy of an original, if such are conceivable. And that, I think, is incompatible with a model which purports to be physiologically *and* linguistically motivated.

Ladd (1992, 1993) and others point to a comparable "degrees of freedom" problem for the tone scaling and tonal space parameters of the AM-framework models of English and Japanese associated with Pierrehumbert and her colleagues, but it surely is a problem for Grønnum's model, too, once one goes beyond carefully scripted lab speech. Indeed, this indeterminacy will be a problem for any analysis by synthesis model that is sophisticated enough to simulate the ways in which tonal space and tone pattern interact in speech but relies exclusively on goodness of f_0 fit as a criterion for choosing among parameter settings. In short, there are very pressing research questions that need to be addressed before we have a good model of tone scaling and its relationship to tonal space control, including the overarching one that Grønnum identifies in her review of Möbius's model: What kind of criteria can be applied to distinguish among models or among different parameter settings within a model?

As noted in Grønnum's review, Möbius defends his choice of framework on the grounds that the tonal space parameters in Fujisaki model are physiologically motivated. The basis for this claim is in Öhman's seminal model of Scandinavian "word and sentence intonation" in which he posited two distinct laryngeal gestures for word accents and sentence-level patterns, and suggested that these could be identified with independent activation of two different parts of the cricothyroid muscle (Öhman 1967: 29–30). Fujisaki (1983: 53–54) follows Öhman to posit the same physiological correlates for the different damping functions that he proposed for the phrase command and the accent command. Work on the control of f_0 in speech has not supported this idea. Neither has it identified evidence of separate "gestures" for tonal space versus tone scaling parameters, because there is no compellingly obvious way to conceptualize the task space. In this respect, the articulation of f_0 is fundamentally different from the articulation of spectral correlates of consonant constrictions. There are some suggestive ideas in work on physiological correlates of tone and pitch accent contrasts, such as Gårding, Fujimura, and Hirose (1970), Erickson (1978, 1993), Erickson *et al.* (1995), Beckman *et al.* (1995), Hallé (1994), and Sugito (2003). There is also research such as Herman (2000) and Epstein (2002), documenting perceptible differences in vowel amplitude and voice quality associated with the final lowering effect. These non-f_0 correlates perhaps could help in conceptualizing the task space for tonal space gestures if examined at the articulatory level, as suggested in Herman, Beckman, and Honda (1996). However, the interactions among laryngeal tension, vocal fold thickness, and pulmonary effort are complex and not completely understood. There looks to be a great deal of basic research yet to be done before physiological evidence can be brought to bear directly on the degrees of freedom question.

Another avenue of attack that may yield more immediately applicable criteria is to develop experimental paradigms for assessing whether native listeners treat tone scaling and tonal space separately, as in Herman (2000) and Gussenhoven and Rietveld (2000). Such experiments might be especially useful if paired with studies designed to pin down the meaning differences associated with minimally contrasting melodic contours where tone scaling or tonal space differences seem to act as a primary cue or as an enhancing secondary cue, as in Hirschberg and

Ward (1992), Grice and Savino (1995), Venditti, Maeda, and van Santen (1998), Caspers (2000), and Lee (2000, 2005). As Gussenhoven (1999) points out, however, this avenue of research requires that we look more closely at the types of linguistic functions that are linked to different formal parameters in different languages, and think carefully about how particular experimental tasks might preclude discovery of the use of some pattern of tones, tonal anchoring, tone scaling, or tonal-space settings for a particular function.

This highlights the fact that we need a better understanding of the range of linguistic functions that can be encoded in spoken language melody and of how these functions are realized in related language varieties as well as in different language families. In the next section, therefore, we will briefly describe some of the functions that have been identified, beginning with the "tonemic" function of constituting a small finite set of meaningless contrasting patterns that can be combined with elements from other sets of meaningless contrasting patterns (consonant constrictions and vowel postures) to build an indefinitely large lexicon.

5 A Taxonomy of Linguistic Functions

5.1 Tonemes and Tonal Morphemes

The basic "tonemic" function is most easily illustrated with utterances and words from a language such as the standard Hong Kong dialect of Cantonese. In this variety of Chinese, most words are monosyllabic (that is, any given *zì* probably is a word), and every syllable is specified for one of the tone patterns exemplified by the contrasting word-forms in (7).

(7) a. [wɐi^{55}] 'power' [wɐi^{33}] 'fear; pleasant' [wɐi^{22}] 'guard'
[wɐi^{35}] 'position' [wɐi^{23}] 'surround' [wɐi^{21}] 'person'
b. [wɐʔt^{5}] 'dense' [wɐʔt^{3}] 'revolve' [wɐʔt^{2}] 'kingfisher'

Figure 15.6 shows example utterances of the six word-forms with sonorant rhymes in (7a) produced as citation-form sentences. The extremely low onset of

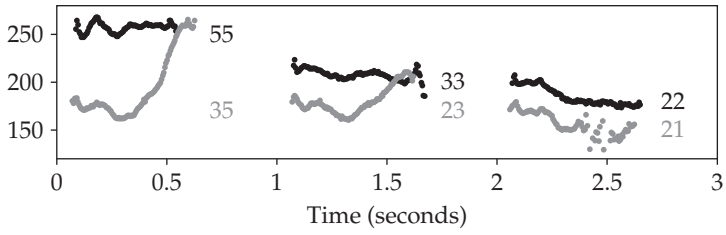


Figure 15.6 Example f_0 contours of citation intonation utterances of the level tone wordforms (black) and contour tone wordforms (gray) listed in (7a) produced by an adult female native speaker of Hong Kong Cantonese. The utterances are from Wong, Chan, and Beckman (2005).

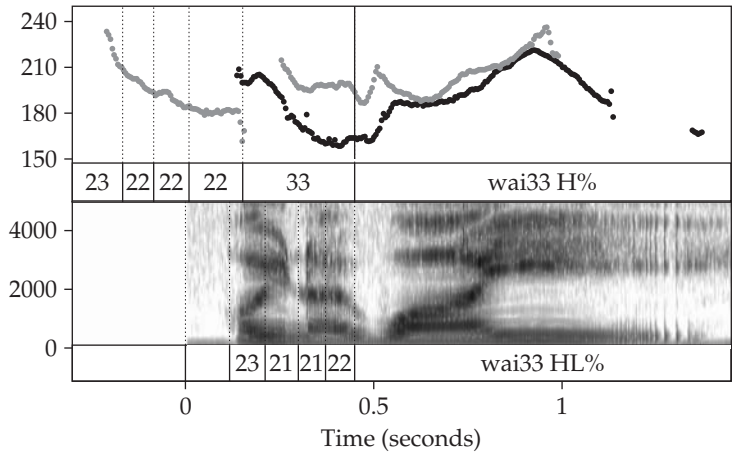


Figure 15.7 Spectrogram and f_0 contour (black) for utterance of the sentence [ɔ^{23}jyn^{21}loi^{21}hɐi^{22}wɐi^{33}]+HL% 'Oh I get it! The word was <fear>.' with overlaid f_0 contour (gray) for utterance of [keoi^{23}jɐu^{22}wɒ:^{22}hɐi^{22}fɒ:i^{33}wɐi^{33}]+H% 'She said then that the word was <pleasant>?!' produced by the speaker who produced the utterances in Figure 15.6. The f_0 contours for the two utterances are aligned at the onset times of the final homophonous words [wɐi^{33}] (solid cursor). The utterances are from Wong, Chan, and Beckman (2005).

the toneme that is transcribed with [^{35}] reflects a sound change in progress in the Hong Kong standard dialect (see So 1996, who transcribes it as [^{25}], and reviews the literature on this and other recent tone changes and merges). The black and gray f_0 tracks in Figure 15.7 illustrate how the mid-level tone of the homophones meaning 'fear' and 'pleasant' is realized in two other intonational contexts. The morpheme just before [wɐi^{33}] 'pleasant' in the utterance plotted with gray in that figure also has this same mid-level tone. The pitch perturbation at the syllable boundary is a juncture-marking creaky voice quality that sets off and emphasizes the final word, which is as long as the total duration of the first four morphemes of the utterance. All of the earlier morphemes in this utterance, as well as in the utterance plotted with the black line, are shortened by a process that Wong (2006) calls "syllable fusion." When morphemes are conjoined into compound words or frequently uttered phrases, speakers can signal the particularly close juncture by weakening or deleting medial consonants and merging the two syllables' vowels. Except in the most extreme cases, however, the percept of each syllable's tone specification is preserved to maintain the syllable count. Thus in this variety of Cantonese the tone specifications are contrastive properties of syllables fully on par with such properties as the palatalized offglide in the rhymes in (7a) as opposed to the glottalized plosive coda in the rhymes in (7b). The typical shapes of words in combination with the extremely "isolating" or "analytic" nature of the grammar drives a robust segmentation of the melody into syllable-anchored tone units.

The pitch patterns on the final syllable in Figure 15.7 illustrate another way in which tones can function in lexical contrast. The [wɐi^{33}] syllable in each of these

utterances is prolonged to be three or five times the average length of syllables earlier in the utterance. This prolongation leaves room for the realization of one or two more tone targets that are transcribed using the notational conventions described earlier for the transcriptions in (1)–(6). This "code-switching" between transcription systems follows the C-ToBI conventions proposed by Wong, Chan and Beckman (2005) to clearly distinguish the morphemic function of the tones transcribed as H% and HL% in these utterances from the tonemic function of the tones transcribed as [^{23}], [^{22}], [^{21}], [^{33}], and so on in (7). The meanings of these two morphemes H% and HL% are reflected in the glosses. The H% at the end of the utterance in gray makes it an incredulous echo question as indicated by the '?!' at the end of the gloss, whereas the HL% at the end of the utterance in black imparts the sense of discovery or sudden realization glossed by the 'Oh I get it!'

Cantonese has an extremely rich set of pragmatic morphemes like these final boundary tones. Many of these sentence particles are composed of vowel and consonant phonemes as well as the tonemes, but several of them are just the toneme affixed to the final content word, as illustrated here. The H% boundary tone in the Beijinghua utterance in Figure 15.4b, similarly, is one of two tonal morphemes among the 28 sentence particles that Chao (1968) counts. The count for any of the Mandarin dialects is somewhat easier, since the non-tonal components of the sentence-final particles of Mandarin are analyzed as being neutral tone and combinations of particles are never counted separately from the particles that are simple syllables or simply tonal. By contrast, counts for Cantonese range widely (as many as 206 by Yau's 1980 count), depending on whether polysyllabic sequences, monosyllabic particles that are potentially fused forms of polysyllabic sequences, and other complex forms are counted separately. For example, Law's (1990) count of between 35 and 40 includes sets that are traditionally described as being minimally differentiated by tone, such as the minimal pair [tse^{55}] and [tseʔk^{5}] studied by Chan (1998). Fung (2000) suggests that Cantonese sentence particles such as these can be grouped into a much smaller number of "families" of phonologically related particles that have a common core meaning. That is, she proposes that the meanings of [tse^{55}] and [tseʔk^{5}], for example, can be analyzed in terms of the core meaning of the [ts] family in combination with the meanings of the tones (which correspond to the tonal morphemes transcribed in C-ToBI as H% versus -%). Sybesma and Li (2007) analyze Fung's families in more detail, and propose that each of the 40 most common sentence particles is composed of three parts: (i) an onset morpheme that is either the default null (glottal stop) initial or one of the fully-specified consonants [h, g, l, m, l~n, ts], (ii) a nucleus morpheme that is either the default vowel [e:] or one of [ɒ:, ɔ:], and (iii) a tonal morpheme that is either the default [^{3}] (tagged as a protracted neutral target :% in C-ToBI), [ʔ^{4}] (tagged as -% in C-ToBI), [^{5}] (H%), or [^{1}] (L%). By this analysis, then, the HL% transcribed for the utterance plotted in black in Figure 15.7 might be a compound of Sybesma and Li's tonal morphemes [^{5}] and [^{1}], or it might be the tonal morpheme corresponding to [^{51}], which as a toneme has merged with [^{5}] in the Hong Kong dialect (see So 1996, among others).

The difficulty of counting the number of Cantonese sentence particles as compared to the ease of counting the nine Cantonese lexical tonemes in (7) is

noteworthy. It may reflect the elusiveness of pragmatic "meaning," which is difficult to paraphrase outside the specific contexts where a pragmatic morpheme is appropriate, as compared to the stark difference in referential meaning that lets us identify the polysemous nature of the word-form [wɒi^{33}]. It also may speak to a more basic difference between tonemes and tonal morphemes that stems from the design principle of duality of patterning (Hockett 1960) – that is the principle that the lexicon of any human language is a self-diversifying system in which a small number of discretely different elements can be combined to make a large number of potentially extremely complex morphemes without losing their discrete distinctiveness (Goldstein and Fowler 2003). Consider the analogous difference for vowel segments. As Hyman (this volume: Chapter 7, Section 2) points out, it is relatively easy to count the number of tonemes in a language such as Dadibi, Nupe, Chatino, Kam, Putonghua, or Cantonese in the same way that it is relatively easy to count the number of vowel phonemes in these languages. And it is relatively easy to recognize that the tones in the second and third syllables of the first utterance in Figure 15.7 are the same toneme and that both are different from the toneme on the first syllable, just as it is easy to recognize that the vowels in the two syllables of the English compound *A-frame* are the same vowel phoneme but that the vowels in the two syllables of *A-team* or *AWOL* are different vowel phonemes. It is harder to say whether the vowel [e:] in the first syllable of each of these words is the same morpheme, or even whether [e:] constitutes a morpheme in *AWOL* in the way that the [e:] in the first syllable of *A-frame* obviously constitutes the first morpheme of a compound word.

These two sources of difficulty have long confounded the analysis of the tonal morphemes of English. Is the tune in the second phrase of Figure 15.3 a sequence of four tonal morphemes, as suggested in the transcription in (2), which follows the analysis in Pierrehumbert and Hirschberg (1990)? Or is it two tone morphemes H*L H*LH to which a linking rule has applied to anchor the L of the first morpheme to the second stressed syllable, as proposed by Gussenhoven (1984)? What kinds of experiments can we use to differentiate between these two morphological analyses? Ladd (2008: Chapter 4) gives an insightful description of the difficulties for English and a few other related languages, as well as a review of arguments advanced by proponents of different analyses and of the relevant experimental studies.

5.2 Prosodic Grouping

In describing the "tonemic" function using Cantonese examples, we emphasized the monosyllabic word shapes and isolating morphology of the language, because the more general function of lexical contrast will be realized using very different segmentation and anchoring parameters in a language where words are polysyllabic or the grammar is of a more "agglutinative" or "synthetic" nature. For example, every modern Chinese dialect has a system of lexical tone contrasts that is a reflex of the same original tone categories that give rise to the Cantonese tonemes in (7), but in a Wu dialect, the tone pattern that corresponds to a toneme

of a Cantonese word typically will not be realized in the same way on the cognate Wu morpheme. Words are typically at least two syllables, and very productive morphological processes (typically called "tone sandhi" – see Chen 2000) insure that just one toneme is specified for each compound word or phrasal construction in an utterance, as illustrated in (8).

The Shanghai examples in (8) are from Zee and Maddieson's (1980) study, and the schematics are based on the f_0 tracks they show. The compound nouns in (8c) and (8d) are derived from the sets of four *zì* in (8a) and (8b), respectively. These examples illustrate the tone sandhi processes that relate the patterns of derived words to the citation-form tone patterns of the *zì* from which they are derived. The most general description is what Chan and Ren (1989) call "Pattern Extension"; the underlying toneme of the first *zì* is the only one realized, and its component tones are extended to cover the whole word or phrase, as in (8c). All of the Wu dialects use some variant of this Pattern Extension process, although details such as the typical phonological anchoring pattern may differ across tone types and across dialects. For example, Zee and Maddieson analyze the abrupt fall in (8d) in terms of a constituent-final tone that they posit for all such compound words, whatever the initial toneme, but they do not discuss the early anchoring point for the tone in some cases, such as (8d).

(8) a. [ɕin^{51}] 'new' [vəŋ^{14}] 'to hear' [t͡ɕi^{34}] 'to record' [t͡se^{51}] 'person'

HL LH MH HL

b. [t͡ɕiʔ^{5}] 'to unite' [h^{w}əŋ^{51}] 'matrimony' [tsəŋ^{34}] 'proof' [sɪ^{51}] 'book'

H HL MH HL

c. [ɕin.vəŋ.t͡ɕi.tse] 'reporter'

H L

d. [t͡ɕi.h^{w}əŋ.tsəŋ.sɪ^{51}] 'marriage license'

H L–

Kennedy's (1953) description of a very similar abrupt fall in Tangxi compounds that have initial syllables with checked tone rhymes suggests an alternative analysis in which the abrupt fall is the realization over longer material of the creaky voice register that characterizes the checked tone. More recent work by Chen (2008) suggests that both analyses may be supported for variant realizations of longer compounds for at least some younger speakers. The cross-dialect differences in anchoring point can be appreciated by comparing the Shanghai falling-tone example in (8c) to the three Wuxi falling-tone examples in (9). The four examples in (10) are an alternative pattern for Wuxi compounds that begin with a falling-tone *zì*.

(9) a. [sɛ] 'three' b. [sɛ.ŋie] '3 years'

H L H HL

c. [sɛ.dʌɯ.mɔʔ.dõ] '3 big wooden tubs'

H HL

(10) [fi] 'fly' [fi.tɕi] 'airplane' [fi.tɕi.p^{h}iʌ] 'air ticket' [fi.tɕi.p^{h}iʌ.tɕia] 'airfare'

HL L LH L LH L LH

The transcriptions and schematics in (9) and (10) are based on the descriptions and f_0 tracks in Chan and Ren's (1989) account of the history of two different morphological processes that they identify in this dialect.[6] They describe the Pattern Extension process in Wuxi as typically applying to number+classifier expressions, as in (9), and also to reduplicated verbs, verbs with resultative or directional complements, and reduplicated nouns in child-directed speech. The Wuxi "Pattern Substitution" process in (10), by contrast, is typically applied to verb phrases with direct objects, to reduplicated nouns in the adult lexicon, as well as to the very productive compound word formation process illustrated in (10), where [fi^{1}.tɕi^{14}] is 'fly machine', [fi^{1}.tɕi.^{1}p^{h}iʌ^{14}] is 'fly machine ticket', and [fi^{1}.tɕi.^{1}p^{h}iʌ^{1}.tɕia^{14}] is 'fly machine ticket price'. Chan and Ren relate these two Wuxi processes to a contrast that Kennedy (1953) describes for the Tangxi dialect, where the morphosyntactic difference is clearer. When the two Tangxi processes apply in combining the morphemes [tsɔ^{51}] 'to fry' and [vɛ^{24}] 'rice' the Pattern Extension process yields the compound noun [tsɔ^{5}.vɛ^{1}] 'fried rice' whereas the second type of process yields the verb phrase [tsɔ.vɛ^{24}] 'to fry rice'.

Despite the differences across the examples in (8)–(10), however, the function is essentially the same. The toneme specification is a property of the constituent as a whole, and the boundaries between successive constituents are marked by a transition to the next lexically contrastive tone pattern. The contrasting melodic contours, then, effectively group strings of syllables into coherent prosodic constituents (tone sandhi groups) that align to constituents or domains specified by other parts of the grammar. When utterances are short and decontextualized, as in the utterances examined in Zee and Maddieson (1980) and Chan and Ren (1989), the domains are described in terms of morphosyntactic relationships. When utterances are longer or produced in more elaborated discourse contexts, other types of relationship, such as the articulation of an utterance into topic and focus or given and new, come into play, as discussed by Selkirk and Shen (1990) among many others.

This same function of prosodic grouping is invoked by Carter's (1974) description of the "tone-phrasing system of Kongo" and by our description of the distribution of L% and H– tones in Japanese in Section 4.1 above. It also is a critical element in Halliday's (1967: 9) description of English utterances as "an unbroken succession of tone groups each of which selects one or another of the five tones" as well as in Pierrehumbert's (1980: 19) definition of the "tune" in English as "the melody for the intonation phrase." As should be obvious from this list of languages, as well as from the differences between Cantonese and Shanghai, the ways in which melody is harnessed for the function of prosodic grouping are orthogonal to the ways in which melody is harnessed for the function of lexical contrast. Cantonese and Shanghai have inherited a cognate set of toneme categories from their common ancestor language, but Cantonese does not have any morphological process like these "tone sandhi" processes in Shanghai and the other Wu dialects and instead uses the consonant- and vowel-focused process of syllable fusion. Thus, the surface melodies of cognate compound words and phrases make for very different tone groups in the two languages. The modern Japanese dialects offer the complementary evidence, for a double dissociation. Although the accentual phrase melodies of Japanese mark off analogous tonally delimited prosodic phrases in very similar ways in the Tokyo and Osaka dialects, the tone at the AP boundary in Tokyo is invariantly L, whereas Osaka preserves an older tonemic contrast between %L-beginning and %H-beginning words.

5.3 Metrical Prominence

In accounting for the melodic differences between the disyllabic compound noun [tsɔ^{5}.vɛ^{1}] 'fried rice' and the verb phrase [tsɔ.vɛ^{24}] 'to fry rice' in Tangxi, Kennedy (1953) talked about the prosodic grouping function that the two patterns have in common, but he also described differences in the "stress pattern," with the compound-noun pattern having "louder stress" on the first syllable and the verb phrase having it on the second. A related segmental difference is specific to the checked tone; the glottal coda in morphemes such as [baʔ^{3}] 'white' or [tɕ^{h}ʊk^{5}] 'to drink' can trigger gemination of a following syllable onset in the compound-noun pattern but not in the verb phrase pattern, as in [bas^{3}.sɛ^{51}] 'white water' versus [tɕ^{h}ʊ.t^{h}saŋ^{33}] 'to drink soup', an interaction that is reminiscent of the stress conditions on Raddoppiamento Sintattico in many varieties of Italian (see, e.g. D'Imperio and Gili Fivela 2003).

In other Wu dialects, also, the syllable in the tone sandhi group that is associated with the toneme bears segmental hallmarks that are associated with phrasal or lexical stress in other languages. For example, Zee (1990) documents a process of vowel lenition in Shanghai whereby the high vowels of the language [i, ɹ, u]^{7} can be devoiced or deleted in certain environments. This is essentially the same process as the "syncope" that Cedergren and Simoneau (1985) describe for [i, y, u] in Quebec French, and the vowel "reduction" that Dauer (1980) describes for [i, u], which Arvaniti (1994) uses as a metric of stress on pretonic syllables. As in these other two languages, devoicing in Shanghai is a variable process that depends

on speech rate as well as on the identity of the neighboring consonants. It is also tonally conditioned. In Quebec French, syncope never affects vowels in the final syllable of the constituent that Cedergren *et al.* (1990) define as the domain of final pitch accent. This is the constituent that Jun and Fougeron (2002) call the accentual phrase, highlighting the demarcative function of the obligatory final pitch accent and the optional initial rise. In Greek, similarly, devoicing does not occur on the stressed syllable in a word – that is the syllable which is associated with one of the tonal morphemes of the utterance melody. In Shanghai, too, devoicing never affects the first syllable in the tone sandhi group – that is the syllable to which the toneme is associated phonologically. Chao (1968: 31 and 141) also notes high-vowel devoicing in neutral tone syllables in the Beijing dialect of Putonghua, a relationship that he describes in Chao (1932: 129) in terms of the notion "stress accent" or "tonic stress":

> Stress-accent does not play any important role in most Chinese dialects. But in a few dialects, including that of Peiping, tonic stress plays such an important part that unstressed syllables not only tend to have their vowels obscured, but also lose their proper tones, and acquire a level, usually short tone, the pitch being determined by the preceding syllable.

Thus, in all four of these languages, the property of being eligible to bear an associated toneme or tonal morpheme prohibits application of a process that weakens or deletes vowels. This is true both of Beijinghua and Greek, where the location of this "tonic stress" is not predictable from the prosodic grouping into words and phrases, and of Shanghai and French, where the fixed position of the "tonic stress" serves to demarcate the tone sandhi group or accentual phrase.

By contrast, high-vowel devoicing is not constrained by the tone pattern of the accentual phrase in either Japanese or Korean. Maekawa and Kikuchi (2005) observe devoiced vowels in the Corpus of Spontaneous Japanese in many syllables that are aligned to the phrasal H– in unaccented phrases or associated to the H of the H+L lexical pitch accent. Jun and Beckman (1994) likewise document pervasive high-vowel devoicing in a corpus of enacted lab speech dialogues in Korean, both in syllables that are AP-medial and in syllables that are associated to the LH or HH sequence that marks the beginning of the AP.

This reduction of vowels in the first syllable of the AP in these two languages contrasts with other segmental effects in this position. In Japanese, for example, older speakers who produce the nasal allophone of [g] in AP-medial positions (such as in the *-ga* particle in the *Yamano-wa oyoideru-ga* clause in Figure 15.1b) do not produce [ŋ] in AP-initial position. Keating *et al.* (2003) and others show that the beginning of the AP in Korean also is a position marked by "initial strengthening" of the consonant. The consonantal effects are more in line with the segmental effects of metrically strong position in languages with tonic stress. In Shanghai, for example, voiced stops are voiceless with breathy voice releases when they are onsets of syllables at the beginning of a sandhi tone group, but are voiced with short closures in tone-group medial position where the syllable does not bear a tone specification (Cao and Maddieson 1992). That is, they show the

same allophonic patterns with respect to the tone sandhi group that Jun (1993, 1996) and others document for the Korean lax stops in AP-initial versus AP-medial position. There are similar effects in syllables with tonic stress versus syllables with neutral tone in the Beijing dialect. For example, the voicing of the [ts] in the second syllable of *fàng zài* in Figure 15.4a is a cue to the neutral tone status of the syllable. One possible way to characterize the different treatments of the vowel in Japanese and Korean as compared to these other four languages, then, is to say that the vowels are less important than the consonants in defining the syllables and the rhythms of anchoring points for tones in these two languages. Another way to characterize the difference is to say that Japanese and Korean emphasize edges at all levels of metrical structure, from the consonant-focused definition of the syllable to the primarily demarcative use of tonemes and tonal morphemes, whereas Beijing Mandarin, like English, emphasizes heads.

This difference has ramifications for the realization of focal prominence. In Korean and Japanese, focal prominence is realized primarily by an expansion of the tonal space to enhance the demarcative rise at the beginning of the first AP coupled with a post-focal erasure of AP boundaries (see, e.g. Venditti, Jun, and Beckman 1996). Other prominence-enhancing mechanisms include the choice of IP-final boundary tones such as the H% tone of Tokyo Japanese (see, e.g. Venditti, Maekaewa, and Beckman 2008). In Beijing Mandarin, English, and Swedish, by contrast, focal prominence instead singles out a syllable with tonic stress and then either reduces or deletes the tones associated to following stressed syllables (see Jin 1996; Xu and Wang 2001, among many others, for Mandarin, Chapter 6 of Ladd 2008, for a review of the literature on English, and Bruce 1977, 1982, among many others, for Swedish). Chapter 7 of Ladd (2008) gives a particularly insightful discussion of this difference between edge-focus and head-focus strategies. He also suggests a common underlying unity. The syllable with tonic stress in languages such as English and Swedish plays a culminative role in marking words and larger morphosyntactic constituents, as illustrated by the tone pattern that marks the compound word in (5). The word that is the domain of the focal H– in Stockholm Swedish, similarly, plays a culminative role in marking intonation phrases and their alignment with the domains of focus in the information structure of the sentence. Pierrehumbert (1980) posits a similar "phrase accent" for English, as in the transcription in (2). Gussenhoven (1984), by contrast, treats the rise-fall-rise over *messenger* and *ball* as a single H*LH tonal morpheme. By either analysis, however, the word that contains the stressed syllable to which the H* tone is associated plays a similarly culminative role vis-à-vis the focus domain in English. Ladd (2008: 278) suggests that both the culminative function and the demarcative function can be viewed as ways of identifying levels of grouping in the metrical hierarchy of a language:

> If . . . we see prosodic phrasing as the ultimate basis of sentence stress, we may see that the correct way to pose questions about universals of the prosody-focus link is not 'Why is the main accent in this sentence on word X rather than on word Y?' but rather 'Why is this sentence divided up into phrases the way it is?'

A further advantage of thinking of "sentence stress" in this way is that focus and other aspects of information structure at the level of the sentence then become the local expression of the same types of discourse structure relationships that are encoded in such effects as final lowering, as discussed in Nakatani (1997) and Venditti (2000).

6 Defining Tone

As the preceding section should make clear, it does not seem very useful to talk about "stress" as if it were an autosegmental content feature, on a par with tone features, manner features, and place features. Rather, stress is better treated as a syntagmatic property of nodes in the prosodic tree, like the property "syllabic" (which is another way of naming the autosegment that stands as the head of some σ). In many languages, stressed nodes in the prosodic tree are defined first and foremost by the licensing of tonal autosegments. In some language varieties, such as Hong Kong Cantonese, syllabic nodes also are so defined (cf. Wong *et al.* 2005, and the discussion of "syllable fusion" above). In many languages, tones also are stereotypically used to mark the edges of prosodic constituents above the prosodic word, a function that is less commonly associated with vowel or consonant features. Tone is remarkably versatile in the roles that it plays in realizing and interpreting the prosodic trees of spoken languages. As Hyman (this volume: Chapter 7, Section 5 (emphasis in the original)) puts it, "*Tone can do everything that segmental and metrical phonology can do, but the reverse is not true.*"

This versatility raises again the question with which we started this chapter: How can we define intonation in a way that delimits the scope of this chapter from the scope of the chapter on tone? Or focusing the question the other way around: What is tone, and how does it differ from intonation?

In his chapter on tone in this volume, Hyman (this volume: Chapter 7) poses the first part of this question, but then replaces it with a question about language types: "How do we know if a language has tone?" In answer, he repeats his earlier "working definition of tone" (Hyman 2006b: 229), which is adapted from Welmers's (1959) definition: "A language with tone is one in which an indication of pitch enters into the lexical realisation of at least some morphemes." He rejects a distinction between "pitch accent" and "tone" (corresponding to Voorhoeve's distinction between "restricted" and "unrestricted" tone systems described above). In his earlier paper, Hyman points out that systems cited as examples of the "pitch accent" type are a varied lot, including languages as different as Tokyo Japanese (where "accent" does not imply metrical prominence and the majority of native Yamato-stratum words are unaccented) and Stockholm Swedish (where every word has at least one syllable with tonic stress and a compound word has exactly two). While he rejects the idea of this third type, however, he maintains that a tenable distinction can still be made between a "tone language" prototype and a "stress-accent language" prototype. His criteria for setting up this distinction require that he treat stress as a "suprasegmental" property on a par with H tone,

rather than as a structural property on a par with syllabicity. That is, he proposes that the prototype stress-accent system is one in which "every word has at least one stress accent" and "the stress-bearing unit is necessarily the syllable."

We cite Hyman (2006b) here because this paper is very representative of a widely-held assumption: that there are fundamental prosodic differences among spoken languages which naturally fall out from the difference between using tone "to make semantic distinctions" and using it "to add functional meaning." This assumption is at the heart of nearly every typology of tone and intonation systems. The difference that is deemed critical in these typologies is a distinction between the tonemic function of lexical contrast and everything else – between languages such as Cantonese, where many of the tones in the melody of a typical utterance are tonemes that combine productively with the consonant and vowel phonemes of the language to make a large and expandable set of morphemes, and languages such as English, where the tones are pragmatic morphemes chosen from a small and relatively closed set. This is a useful distinction, because it predicts that there would be sharp differences in native speakers' and linguists' metalinguistic awareness of the tone count, as suggested in Section 5.1. However, contra Hyman (2006b), we do not see that it correlates neatly with all of the other distinctions that could be made on the basis of the functions outlined in Sections 5.2 and 5.3. That is, we can appreciate the difference in ease of counting tones in Putonghua versus English that falls out from the fact that a L+H that is anchored to a stressed syllable in Putonghua is a toneme whereas a L+H* that is anchored to a stressed syllable in English is a pragmatic morpheme. But this difference does not change the fact that these two languages are far more like each other in many other respects than either is to a language such as Japanese. There is no useful classification of prosodic types that falls out from the classification of languages in terms of the tonemic function alone.

Hyman begins his chapter on tone by saying that, "Except for a brief period in the late 1970s and early 1980s, tone has generally fallen outside the central concerns of theoretical phonology" (Section 2). He ends it by calling for renewed attention to questions about "the INTERDEPENDENCY of tone with other features" – questions such as "What is or can be a TBU?" and "What is the correct set of tonal features?" (Section 5). He concludes by saying that "the above set of questions . . . may even contain a misunderstanding that we still have either about tone, or more likely, about phonology in general." (Section 5)

As we have tried to make clear in this chapter on intonation, we think that, contrary to Hyman's assessment of the last three decades, tone *qua* intonation has been very central to major developments in theoretical phonology, and that work on intonational phonology has addressed the questions about interdependencies and proposed answers that suggest that the questions do indeed "contain a misunderstanding" about the nature of tone. That is, the work that we reviewed in Section 4 suggests to us that the presuppositions of these questions constitute the tonal counterparts to the assumptions of "segmental idealization" and "universal categorization" in segmental phonology, as discussed by Ladd (this volume). In short, we suspect that defining tone (and delimiting intonation from it) in terms

of the tonemic function alone may have delayed progress in our understanding of tone. It may have obscured the true diversity of ways in which speakers of different languages use pitch variation to structure their words, utterances, and larger spoken discourses. We close by reminding the reader that the languages for which we have a good solid description of the system of lexical tone contrasts vastly outnumber the languages for which we have even a cursory description of the tone patterns that are associated with other levels of the grammar. Until that gap in coverage is filled in a bit more, any delimitation of tone from intonation based on a classification of language types seems premature.

NOTES

1 Sections 3–5 of this chapter are very nearly identical to sections 2–4 of our chapter on "Tone and intonation" in the second edition of the *Handbook of Phonetic Sciences* (Hardcastle, Laver, and Gibbon, eds., 2010). The material covered in these sections is reused here with the concurrence of the editors of this volume, since it supports the different mandate of this chapter (which was to review the contributions that intonational phonology has made to phonological theory) just as well as it supported the mandate of the other chapter (which was to lay out the bases of a taxonomy of forms and functions of linguistically significant pitch variation).

2 Chao's tone letters locate notes and glissando turning points in the local tonal space in terms of five points numbered from 1 for the bottom to 5 for the top of the tone space.

3 The traditional use of the term "accent" both for pragmatic tonal morphemes in English and for lexically-specified tone patterns in Japanese is the source of frequent confusion among scholars of both languages. Further confusion is caused by the fact that the Japanese word *akusento* which "accent" translates here refers to the entire configuration of tones for the level of prosodic grouping that is called the accentual phrase, including both the lexically-specified pitch fall at the designated syllable and the "post-lexically" specified pitch rise at the AP beginning. See Venditti, Maekawa, and Beckman (2008) for an explication of the differences between the two phenomena.

4 See also (67) in Hyman (this volume), where the fully-specified account is assumed and the surface tone string transcribed with Africanist tonal diacritics.

5 While we focus on the tonal aspects of the definition here, each of these levels of prosodic grouping is also marked by segmental effects such as differing degrees of "initial strengthening" and "phrase-final lengthening (see review of this approach to these phenomena in Fletcher 2010).

6 Hyman (this volume: Chapter 7, Section 6) analyzes these patterns in term of tone-spreading to result in a fully-specified surface tone pattern, but this is not the only possible analysis (see Yip 1989).

7 Following Zee and Maddieson (1980), we transcribe the apical vowel here and in the examples with [ɹ] rather than with the non-IPA symbol used in the Sinological literature.

16 Dependency-based Phonologies

HARRY VAN DER HULST

1 Introduction

This chapter presents an approach to phonological structure which places head–dependency relations as central organizing relations. This leading idea originates in two models called *dependency phonology* and *government phonology*, which today occur in a number of varieties.[1] In this chapter, I present a synthesis of the leading ideas that are shared by all these models, which, for convenience, I subsume under the name *dependency-based phonologies*. In Section 2, I discuss the notion *head*, while Section 3 contains some remarks on the organization of the grammar and the place of phonology in it. The use of head–dependent relations in segmental structure and syllabic structure is dealt with in Section 4 and Section 5, respectively. In this chapter I will not discuss head–dependent relations at higher prosodic levels (such as the foot). In Section 6 I discuss the use of so-called *empty-head rhymes*, which provides the necessary background for a discussion of relations that can be invoked to control the distribution of empty syllabic, and also branching, units which can be interpreted as head-dependency relations of a special kind. In Section 8, I briefly discuss some hallmarks of currently prevailing varieties of dependency-based phonological theories. Section 9 offers some concluding remarks.

2 The Notion "Head" and Related Notions

The theories discussed in this chapter all make use of organizing relationships between *heads* and *dependents*. The relationship between the head and its dependent

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

is called a *head-dependency* relation or a relation of *government* (where the head is said to *govern* its dependents). Some of these relations define configurations that closely correspond to constituent structures, while in other cases they are used to augment constituency relations. Although dependency approaches to sentence structure have a long tradition (see Fraser 2005), specific dependency formalisms have been developed as alternatives to constituency-based grammars (Hays 1964; Robinson 1970), whereas in other cases notions such as *head* and *dependent* have been added to the daughter nodes of constituents (Chomsky 1970). In addition, in some of the frameworks discussed here, there are further relations between units that do not necessarily replace or augment constituency, but which are also said to involve a head–dependent relationship. The difference between what I will call *structural* and *non-structural* relationships can be compared to the difference between relationships between syntactic heads and their modifiers (where the terms *complement* and *specifier* are also used) and, for example, the relationship between a constituent and an anaphor that this constituent is somehow co-linked with. Indeed, the approaches discussed in this chapter share the assumption (which Anderson 1986, 1992a, 2004, 2006 terms the *Structural Analogy Assumption*) that different modules of the grammar employ identical relationships. Other notions from syntax play a role in the discussion here, including *visibility*.

In phonology, in (morpho)syntax and in semantics, the terms *head* and *dependent* have been used in a variety of approaches, and in several works the question is raised as to what a head *is* (e.g. Zwicky 1985a; Corbett, Fraser and McGlashan 1993). It is one thing to define what is understood to be a head within one specific module of the grammar, for example, in syntax, but if the term *head* is used in various (or even all) components of grammar, the further question arises as to what extent this cross-modular use of a single term is justified. To find out whether heads in phonology and syntax belong to the same type of entity, or at least are both species of one genus, one must not take a specific definition that has been proposed in one domain (e.g. syntax) only to claim that this definition does not fully match entities that are called *heads* in another domain (e.g. phonology). We need to compare the properties of alleged heads in both (or rather, all) domains, and focus on the question as to what the true commonalities are. If there are none, it might be advisable to adopt different terms for each domain, which is not the same as saying that any given domain has a natural right to any given term. If there are commonalities, they may need to be stated in general terms that allow the traits found in each specific domain to be regarded as possible instantiations of these terms. In addition, it may be that, due to independent differences between modules, heads (and dependents) have modality-specific connotations. What is said here about the term *head* applies equally well to terms such as *government* or *licensing*.

The most general characteristic of perhaps all uses of the term *head* (and *dependent*) seems to be that all relationships (whether structural, i.e. between sisters in constituency-based notations, or non-structural) between a unit called the "head" and a unit called the "dependent" are *asymmetrical*:

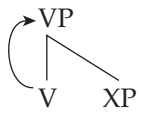
(1) *Asymmetrical*: If A is the head of B, B cannot be the head of A

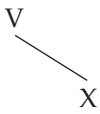
But that in itself says very little. An additional important aspect of head–dependent relationships is that while a head can stand on its own, dependents presuppose the presence of a head. This is sometimes stated as follows: dependents need to be licensed by a head, while the head needs no licensing. This does not exclude that in specific cases dependents are *required*. This latter situation holds when verbs select a complement, or when vowels (such as lax vowels in Germanic languages) require a following consonant (van der Hulst 1985).

To give more content to the notions *head* and *dependent*, we need diagnostics that allow us to identify units as either one or the other. I will single out two such diagnostics: *visibility* and *complexity*. Compared to dependents, heads are *more visible* and (often) more *complex*. These diagnostics are relevant in both syntax and phonology which, in my view, supports the idea that there is a notion of *head* (and *dependent*) that generalizes over these two domains in a meaningful way. We consider these two concepts in order.

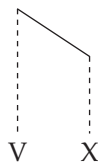
A specific diagnostic of a head can be called *visibility*. Formally, in constituency-based syntax, features of heads are projected to the mother node of a constituent (cf. 2a). If vowel harmony is construed as a relation between syllables (or perhaps rhymes), the featural properties of their head vowels *must* be visible in the syllable (or rhyme) node. By assuming that syllables (or rhymes) are headed by their vowels and projection takes place, vowel harmony can be construed as a local relation between syllable (or rhyme) nodes, as suggested in van der Hulst and van de Weijer (1995).

The relevance of the formal notion of *projection* is dependent on a particular way of representing relationships, namely in the form of constituency (viz. 2a). If a direct dependency presentation as in (2b) is adopted, it follows necessarily that the head properties are the ones that count at higher levels of organization, simply because the head *is* the node that characterizes the whole constituent:

(2) a. 

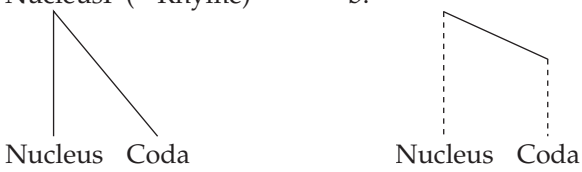
b. 

The dependency notation in (2b) is a notational variant of another notation for dependency trees (cf. Anderson and Ewen 1987):

(3) 

If anything along the lines in (2) and (3) is an adequate way of representing syntactic structure, then, given the analogies between syntactic structure and phonological structure that have just been mentioned, there is no reason to reject a representation of rhymes in terms of, for example, (4):[2]

(4) a. NucleusP (= Rhyme) b.



All these notations are means to express headedness *and* linear precedence, the latter in terms of the linear precedence in a two-dimensional plane, which represents a linear precedence between the units that enter into the dependency relation. If linear precedence is not relevant (either not at all, or not at the relevant level of representation), no slanted (constituent and/or dependency) lines are necessary and a notation such as the following can be used:

(5) Nucleus
|
Coda

Such representations (which are perfectly fine in strict dependency notations) can be used if the linearization of nucleus and coda can be attributed to a different subcomponent of the grammar (see Anderson 1986). Notations of this kind are, in fact, common with reference to segment-*internal* structure in models in which head–dependent relations are postulated between the (co-temporal) features that make up phonological segments. Dependency phonology, indeed, uses dependency relations between monovalent features (here called *elements*).

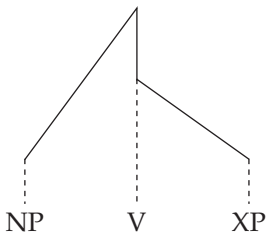
(6) a. |A| b. |I|
|
|I| |A|
[ɛ] [e]

(|A| represents lowness while |I| represents high-frontness.)

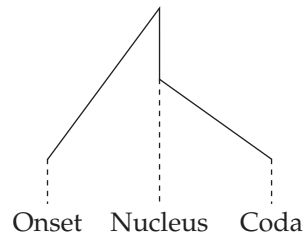
I will discuss segment-internal structure in more detail in Section 4.

Returning to syllable structure, there can be (and often has been) noted a further analogy between the relations between sentences and syllables, which allows us to highlight a specific type of dependency:

(7) a.



b.



Note here that the heads (i.e. V and Nucleus) are represented as heads *of themselves* in order to capture, in a dependency notation, that the relation between heads and following material is more intimate than between head and preceding material (cf. Anderson and Ewen 1987). The dependency graphs in (7) reconstructs the NP-VP division and the onset–rhyme division, respectively.

The structural analogies between sentence structure and syllable structure are striking and it would therefore be counterproductive to approach both types of linguistic units in different structural terms. Indeed, John Anderson has claimed that dependency trees (and thus head–dependent relations) are adequate notational devices both in phonology and in morphosyntax. This point is independent of the kind of grammar formalism that is adopted in both domains and it holds equally well in case one uses constituency-based formalisms. More than any other phonological models, dependency phonology has been founded on the premises that dependency (and thus headedness) is a foundational concept in phonology and indeed in grammar at large.

In any event, whatever formalism is used, heads have a greater visibility than dependents, which is either expressed by appealing to projection or, more directly, by having heads dominate their dependents. Even in the representations in (6), visibility is a diagnostic of heads in that the structure in (6a) is meant to be acknowledged in the phonology as an |A| type of vowel, that is forming a natural class with other |A|-headed vowels, whereas the structure in (6b) represents an |I|-type vowel.

Returning to the analogy between syntax and phonology, Dresher and van der Hulst (1998) have noted that dependency relations, especially at higher phonological levels (i.e. above the syllable) indeed appear to be of a different kind than the relations discussed so far. Whereas the relationships discussed so far (for phrase structure in syntax and for syllable structure in phonology) hold between units *of different kinds*, namely atomic units (such as N, or V) and units that themselves have, or can have, an internal head-dependency structure (i.e. phrases),[3] relationships within the (metrical) foot hold between syllables, that is, entities of the same kind. Dresher and van der Hulst call these two kinds of relationships α-β and α-α relations, respectively. The latter kind of relations may be unique to phonology, having no counterpart in morphosyntax, although this, as always, depends on one's analysis of phenomena such as conjunction or, more specifically, compounding. Additionally, intrasegmental relations (between features or elements) are of the α-α type, but this may equally be true for the features that make up syntactic categories.

However, whatever the reason may be for a modularity difference between phonology (allowing α-α relations at higher levels) and morphosyntax (not allowing α-α relations), this difference is irrelevant to the general claim (central to the dependency approach) that *all* relationships are headed. And, indeed, in all strands of metrical/prosodic phonology, as well as in dependency phonology, it has been assumed that higher levels of organization are headed. This brings us to a second diagnostic trait of heads and dependents, namely (relative) complexity.

A typical example of a complexity asymmetry occurs when accented syllables have specific phonotactic properties such as allowing more syllabic complexities (e.g. syllable closure, branching onset), or allowing contrastive specification of length or tone. Dresher and van der Hulst (1999) identify a difference of this kind as a fundamental asymmetry between heads and dependents. They say that dependents can be less, but not more "marked" than heads. A typical result of this asymmetry, one that Dresher and van der Hulst single out, is that dependents display neutralization of contrast. This does not only play out in terms of syllabic complexity (when, for example, accented syllables allow branching onsets, whereas unaccented syllables do not), but it also affects the content of phonemes. We often see that the array of vowels in the accented syllables is greater than that in unaccented syllables. However, cases of neutralization can involve effects that are apparently at odds with the head-dependent asymmetry, proposed in Dresher and van der Hulst (1997). It is very common to find a process of vowel lengthening which neutralizes the distinction between long and short vowels in *accented* syllables only. In other words, neutralization of contrast may hit both heads or dependents. What this demonstrates is that the desired differentiation between accented and unaccented vowels can be achieved in two ways, both serving the polarization between heads and dependents. While dependents can thus display a greater array of contrast (allowing both long and short vowels, while heads only allow long vowels), the head–dependent asymmetry remains intact since even in that case dependents do not allow greater complexity:

(8) *Complexity constraint*
The maximal complexity of dependents cannot exceed the maximal complexity of heads

This, crucially, forbids a case in which head vowels must be short, while the dependent position allows long vowels.

Asymmetries of this kind imply that the parametric settings for vowel inventories or syllabic inventories are relative to head and dependent positions of the relevant units, which underlies the relevance of the head–dependent distinction for parameter setting (cf. van der Hulst and Ritter 2002).

Dresher and van der Hulst (1998) point out that we also find the notion of projection at higher levels of prosodic organization. Due to their visibility, certain conditions can be imposed on heads which cannot be imposed on dependents because the internal structure of the latter is not visible. This leads to circumstances which are apparently at odds with the principle in (8). If, for example, a condition

is that a unit may be at most binary branching, dependents may get away with being more complex, simply because their internal structure cannot be seen, while heads must adhere to this requirement. I refer to Dresher and van der Hulst (1998) for examples of this phenomenon.

The above supports the claim that head-dependency relations are just as relevant in phonology as they are in syntax and that, moreover, heads in both domains are species of a common genus. Projection or visibility and complexity can both be understood as consequences of what I take to be the fundamental idea behind headedness, which is this: if in a combination [AB], B is the head, we mean to say that the combination as a whole is "a kind of B"; [AB] is a subclass of [B]. We interpret this as meaning that [B] (in the combination [AB]) is in some sense the *central* unit, and that most of the properties of [AB] come from [B], with [A] contributing only some of its properties.

In this section I have used the terms *head* and *dependency*, as they are used in dependency phonology. The term *government* (which underlies the name of the model called *government phonology*) is simply the inverse of dependency. Thus, a head can be said to *govern* its dependent. The term *licensing* is often used interchangeably with *government*. Since a dependent cannot exist without a head, it can be said that a head *licenses* its dependent(s). In Section 6, I suggest using the term *government* for structural head-dependency relations and *licensing* for non-structural (also called *lateral*) relations.

3 Phonotactic Structure, Constraints, and the Organization of Grammar

Like all other phonological models, dependency-based phonology aims at characterizing properties and computations that pertain to the phonological properties of grammatical expressions, that is, simplex units (i.e. morphemes), as well as morphological and syntactic constructions. Specifically, there is a set of elements,[4] partly related in terms of headed relations (cf. 6), which are associated to segmental root nodes, which in turn associate (perhaps via so-called *skeletal positions*[5]) to syllabic nodes, namely *onsets* and *rhymes*, which also display head–dependent relations internally. Additionally, between the onset and rhyme nodes there are headed relations (here called *syntagmatic licensing constraints*) which control the distribution of branching or empty occurrences of these two units, that is, those types of syllabic units that contribute to the complexity of phonological representations (i.e. deviations from the unmarked CVCV . . . sequence):

(9)

Onset and rhyme nodes
:
(Skeletal positions)
:
Root nodes
:
Elements

The combined linkage of elements to a single segmental root node is subject to *cooccurrence constraints*. Likewise, there are *structural constraints* on the number of root nodes or skeletal positions per syllabic unit as well as constraints on the elemental content of these nodes (*paradigmatic* or *positional content constraints*). Additionally, there are *syntagmatic content* constraints that control the content of root nodes with reference to each other (such as vowel harmony); the term *sequential* is sometimes used as an equivalent to "syntagmatic."

(10) Typology of Phonotactic Constraints

a. Complexity of syllabic constituents (Section 3)
   i. Onset obligatory (yes/no)
   ii. Branching (yes/no)
   iii. Empty-headed (yes/no)
b. Licensing of marked syllabic constituents (Section 6)
c. Element cooccurrence constraints
d. Positional content constraints
e. Sequential content constraints[6]

A *phonotactic representation* is only one dimension of a linguistic expression, that is, an expression generated/admitted by the mental grammar of a speaker/listener, the other dimensions (minimally) being, a *semantic representation*, and mediating between these two, a *morpho(syn)tactic representation*. A *morphotactic* representation results from combining or merging (in terms of constituency and/or dependency relations) linguistic units (morphemes, words, phrases) into larger units in accordance with a set of *morphotactic constraints* which guarantee well-formedness. Each such product then also needs to be examined by the grammar for its phonotactic and semantic well-formedness, a procedure formerly called *interpretation*. We can see this procedure as involving *checking* whether (morphotactically well-formed) merge products are also well-formed phonotactically and semantically, employing these terms in a sense similar to their use in current syntactic theory. If this is not the case, and here I will only be concerned with the phonotactic side, an expression is ill-formed. Since phonotactic requirements and semantic requirements are fundamentally different in nature, it is to be expected that many attempts to merge units will not meet the constraints imposed by both subcomponents of the grammar and for that reason, presumably, mental grammars provide computational mechanisms to repair at least some of the phonotactic representations that are delivered by the merge procedure, namely those that turn out to be ill-formed phonotactically. Different models have somewhat different ways of implementing the need for repair rules. As in standard generative phonology (Chomsky and Halle 1968), dependency phonology adopts phonological rules that perform repair operations, being agnostic on the need for extrinsic ordering between these rules.[7] Government phonology builds repair into what is called *phonetic interpretation*.

4 Intrasegmental Dependency

Both dependency phonology and government phonology make use of asymmetrical relations between elements that form the content of root notes and/or skeletal positions. In the former model, this relation is said to be a *head-dependency relation*, whereas the latter model avoids this terminology, seeing the dependent elements as an *operator* (cf. Kaye, Lowenstamm, and Vergnaud 1985). In this section I discuss the motivation for using intrasegmental dependency in combination with monovalent elements.

Dependency phonology did more than introduce the concept of dependency in phonology. A second hallmark of this model was the consistent use of *monovalent* (*unary, single-valued, privative*) *primes*, thus rejecting the binary features stemming from Jakobson, Fant, and Halle (1952) and Chomsky and Halle (1968). From the outset it is crucial to emphasize that the use of monovalent primes does not break with the idea that contrast is a foundational notion in phonology. Logically, a contrast between two segments can be expressed in the following two ways (not considering other possibilities):

(11) a. /p/ [–voice] /b/ [+voice] b. /p/ /b/ [voice]

Generative phonology started out with (11a) and then developed markedness theory and, subsequently, (radical) underspecification theory (Kiparsky 1982b; Archangeli 1984, 1988a; Archangeli and Pulleyblank 1994) in order to be able to express an apparent asymmetry between the two values of almost all features. In the extreme case, one of the values is completely inert, that is, entirely invisible to the phonology. This can be accounted for by adopting the mechanism of underspecification (which had originally been proposed to capture redundancy). But, as Steriade (1995), which offers a general discussion of these issues, points out, a more radical approach is to simply deny the invisible value any theoretical status. However, rather than proceeding by cautiously turning binary feature, into unary feature, one by one (as suggested by Steriade), we should prefer an approach which begins with the assumption that all features are monovalent (Kaye 1988), for the reason that this step can be falsified more easily than the more cautious approach. Dependency phonology took the monovalency approach in the early 1970s and, following this lead, so did government phonology and also particle phonology (Schane 1984, 1995).[8] To date, the viability of this hypothesis is still being tested. One potential problem in falsifying the strong monovalent hypothesis is that an apparent counterexample, which, say, seems to require reference to both values of a traditional binary feature, could be immunized by proposing that there are two monovalent primes which come close to being each other's opposite (cf. Steriade 1995). However, a difference would still exist between [+F] and [–F] and [F]/[G] (G characterizing the same natural class as [–F]) if the latter approach makes crucial use of allowing both [F] and [G] to be part of a single segment which

could not be mimicked by the binary system (if we disallow intrasegmental sequentiality). This (i.e. the idea of combining monovalent elements, even when they would appear to have antagonistic interpretations) is precisely what we see in the dependency models.

In this chapter I will not review any specific proposals for inventories of phonological elements.[9] Rather I will focus on the use of dependency relations in intrasegmental structure, using the three vowel elements that were originally proposed in dependency phonology (Anderson and Jones 1974; Anderson and Ewen 1987: 206; Ewen 1995):

(12) |A| lowness
|U| roundness
|i| frontness

In (12) I have provided articulatory glosses for the elements. Both dependency phonology and government phonology specifically state that the elements are to be understood as primarily acoustic in nature (in agreement with Jakobson, Fant, and Halle 1952), that is, mental acoustic images (Harris and Lindsey 1995; Ingleby and Brockhaus 2002). However, it seems unavoidable to assume that acoustic images must be linked to articulatory motor programs. After all, in production acoustic images need to be realized. Thus, it would seem inevitable that, for each element, we need to know both the acoustic targets and the articulatory plans, neither of which can be completely invariant given that, as we will see, each element corresponds to a variety of implementations which depend on its status as head or dependent.

Intrasegmental dependency relations are used to differentiate, for example, differences in vowel height, as already shown in (6), repeated here for convenience:

(13) a. |A| b. |I|
|
|I| |A|
[ɛ] [e]

(|A| represents 'lowness' while |I| represents 'high-frontness')

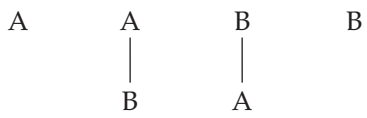
In the representation in (13a) the element |A| is a head, which accounts for its greater salience, that is, its greater contribution to the overall phonetic quality of the vowel, whereas in (13b), the element |I| is more prominent. In terms of projection or visibility, representation, such as, in (13) embody the idea that an |A|-headed vowel structure behaves as a low vowel, that is, on a par with other |A|-headed vowels, including the vowel /a/, which only possesses the element |A|. Conversely, the element |I| in (13a), being a dependent and thus invisible,

could not be used to group the relevant vowel with other, |I|-headed structures. In other words, intrasegmentally, visibility, and salience are properties of intrasegmental heads, as they are of heads in higher levels of organization. Whether intrasegmentally or not, heads can also be more complex (i.e. a head can be itself a compound of more than one element) than dependents (which, then, would have to be at most a single element), is an issue that I will not discuss in this chapter.

In general, there are six advantages of a dependency-based monovalent approach which remain valid even if the details of the structures or their interpretations are modified.[10]

First, by invoking dependency relations, we can strike a balance between systems of phonological primes that allow (in principle unrestricted) use of multivalued features and Jakobsonian systems that only allow binary oppositions. For any given pair of elements, there are four possible configurations:

(14)



The relations in (14) allow a relative (yet restricted) expression of the prominence of any given element and thus the expression of stepwise, quasi-scalar processes, a point that Anderson and Ewen (1987) underscore with reference to accounting for synchronic reflexes of lenition and fortition processes. In fact, we give expression to all three kinds of oppositions recognized by Trubetzkoy (1939). Privative oppositions involve the presence versus absence of a prime. Equipollent oppositions involve the presence of a prime in one member of the opposition and the presence of another prime in the other member of the opposition and, third, as already mentioned, gradual, multi-valued oppositions can be expressed in terms of the way in which a particular component enters in the composition of a class of segments (i.e. as head or dependent).

Second, by replacing binary features with constellations of unary elements, varying in complexity, representations adequately reflect the relative markedness of phonological segments and their properties. In (14), the simple structures (consisting of only one element) are less marked than the categories that are represented in terms of element combinations. Binary notations can only capture such distinctions by augmenting the basic apparatus with an ad hoc system of underspecification.

Third, if we assume (as most phonologists do) that phonological rules can only reflect phonetic events by manipulating phonological units, the set of elements in (12) expresses the claim that languages can have roundness spreading (as an assimilatory process) but not the spreading of non-roundness. If this is empirically correct, the theory in (15) is superior to binary feature systems in which [+round] and [–round] have the same status and are both available for phonological manipulation.

Fourth, given the addition of a head–dependent relation, an impressive reduction in the number of primes can be achieved. In order to characterize major classes and manner distinctions in the feature system of Chomsky and Halle (1968) (or its feature-geometric descendants), one needs many features (such as [voice], [nasal], [lateral], [strident], [continuant], and so on) where unary models use just two single-valued primes, the components |C| and |V| and their interdependencies.

Fifth, the dependency-constellations are constructed in such a way that often-observed affinities between the phonological categories that they represent are formally expressed. For example, just one element |L| is used to express voicing and low tone, and whether one or the other interpretation is relevant is dependent on the head or dependent status of the element. This makes immediate sense of the often-observed connection between these two different phonetic properties belonging to obstruents and vowels, respectively.

To illustrate the two latter points consider the proposal in van der Hulst (1988a, 1988b) to adopt the following interpretations of the three elements in (12) as either heads or dependents:

(15)

	Head	Dependent
|A|	lowness	retracted tongue root
|U|	backness	roundness
|I|	frontness	ATR

This proposal makes explicit that the phonetic interpretations of elements can differ depending on their status as heads of dependents. My goal here is not to discuss the merits or drawbacks of the specific interpretations, but merely to illustrate how the use of dependency can make it possible to account for the relatedness of certain sets of phonetic properties (which are denoted by separate features in other frameworks) by postulating a single element for them. The proposal in (15) has been extended to all elements in the dependency model called Radical CV Phonology (van der Hulst 2005a).

Sixth, it has been argued recently (for example, in Clements 2002) that phonological representations should be minimally specified, that is, with only those feature specifications that are needed in the phonology. Representational minimalism has been a core result of monovalent systems since the inception of dependency phonology and government phonology. The use of monovalent primes largely undercuts discussions about leaving or not leaving out redundant specifications and further contributes to what we could call a "minimal phonology."

Finally, let us briefly compare the use of dependency relations discussed here to usage of this notion in other segment models. In models of feature geometry (as summarized in McCarthy 1988), dependency refers to the fact that certain features are subordinate to other features to indicate that the former are only relevant within the domain of the latter. For example, [lateral] is subordinate to the feature [coronal] to express that only coronals can be distinctively specified

for laterality. Another use of dependency is proposed in Mester (1988), who proposes that a dependency of one feature on another feature accounts for the fact that the former, dependent, feature "gets a free ride" on a process or generalization that is formulated to apply to the dominating feature. Both uses of dependency seem rather different from the use of dependency in dependency phonology and government phonology, but see Ewen (1995) for a more detailed comparison.

5 Intrasyllabic Dependencies

In Section 2 I briefly discussed head-dependency relations at the syllabic level. In this section, I will review the head-dependency structure of segments into syllabic units, and between these units, in more detail.

Dependency-based phonology takes the central units of phonotactic representations to be the *Onset* (O) and *Rhyme* (R). In government phonology it is stated that a regular alternation of O and R is axiomatic (Kaye, Lowenstamm, and Vergnaud 1990). A language that seems to allow syllables without onsets is, in this view, allowed to have Os without segmental content. It *is* possible to defend a different view (following, in this respect, dependency phonology and other OR models) in that we allow a language to have O as an option. Although there is not much difference between these two views (Os can be absent/Os can lack a skeletal point), henceforth I will assume that each R *is* preceded by an O. In Section 6 I will argue that there is an advantage to this view.

I will now, first, discuss the use of dependency relations at the syllabic level in government phonology and then turn to their use in dependency phonology. As we will see, while the former model is based on the notion of constituent structure, the latter is not.

Government phonology[11] (Kaye, Lowenstamm, and Vergnaud 1990) assumes that, depending on a parametric choice, both O and R can branch in a given language (i.e. dominate two skeletal points), or can be "empty-headed," which means that there is only one skeletal point which, however, has no content:[12]

(16) Basic principles and parameters of government phonology

Principles:

a. A phonological representation is a linear arrangement of alternating O(nset) and R(hyme) nodes.
b. Each R universally dominates at least one X-position (that is, the syllable head).

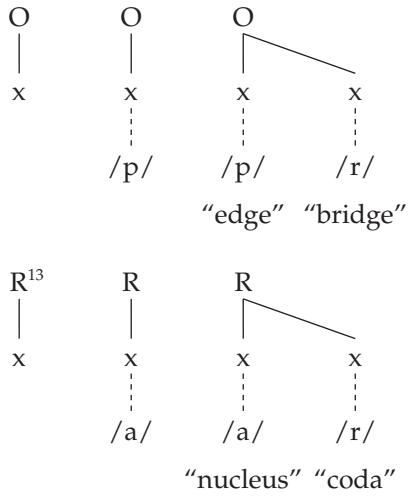
Parameters:

c. Each O dominates at least one X-position ("onset is obligatory").
d. Each O and R may be maximally binary branching.

(The head of an O or R may be empty.)

These five basic principles allow the following six representations:

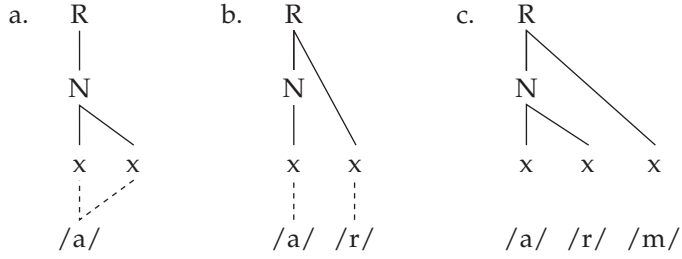
(17) Possible syllabic constituents



Here the terms *nucleus* and *coda* are informal labels for the rhyme head and rhyme dependent position. It is convenient to have an analogous pair of terms for both onset positions and I propose to use *edge* and *bridge*, respectively.

Government phonology has adopted the idea that all constituent structure is strictly binary: a mother node can have at most two daughters. This stipulation has also been put forward in the domain of syntax, for example, in Kayne (1994). Taking this one step further, it has been claimed that onsets and rhymes cannot contain more than two segments. This view excludes treating the sequence /arm/ as a single rhyme.[14] Seemingly in conflict with the latter claim, government phonology (as in Kaye, Lowenstamm, and Verngaud 1990), does make a distinction between Nucleus (18a) and Rhyme (18b) as potential branching constituents:

(18)



However, because it is claimed that ternary rhymes (resulting from combining the branching option in both structures) should be ruled out, an extra principle is stated to the effect that the skeletal point of dependents must be strictly adjacent to that of their heads. In the case of a ternary rhyme (e.g. *arm* in *warm*), this constraint is violated because /m/ is not adjacent to /a/ in the relevant sense.

Van der Hulst and Ritter (1999a), however, propose to deal with this by denying the distinction between nucleus and rhyme, and thus stick to the six representations in (12). In the discussion in Section 7, I will follow this suggestion.

While branching constituents have traditionally been postulated for, for example, prevocalic consonantal sequences of rising sonority (usually taken to form branching onsets), or vowel consonant sequences said to form closed syllables (usually taken to form branching rhymes), a word or two needs to be said about empty-headed constituents, which I will discuss more thoroughly in Section 5. Briefly, these contentless constituents are motivated by cases where an apparent onsetless syllable behaves *as if* it starts with a consonant, or where an apparent non-existing rhyme shows up as a full vowel in related forms. An example of onsetless syllables that behave like syllables starting with a consonant are the so-called h-*aspiré* words in French. These words (such as *hache* [aʃ] 'axe') select the definite article *le* which otherwise only occurs before words that start with a consonant, e.g. *le chat* 'the cat', while words starting with a vowel take *l'*, e.g. *l'ami* 'the friend'.

A crucial aspect of the above representations is that syllabic constituents are *headed*, the head being graphically indicated by the vertical line.[15] In onsets the least sonorous segment (typically or perhaps exclusively an obstruent) is the head, whereas in the rhyme the most sonorous segment claims this privileged status.

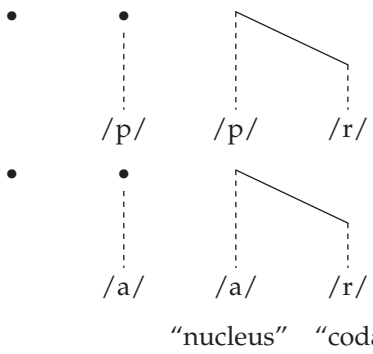
The idea of representing onsets and rhymes as (left-)headed units accounts for the fact that the left-hand member in a complex syllabic unit is the most salient segment *in that unit*. Thus, in complex onsets like /pl/, the /p/ is the optimal (i.e. *least* sonorant unit), whereas in a complex rhyme /am/, the vowel /a/ is the optimal (i.e. *most* sonorant) unit. The visibility of heads is evident from the fact that processes referring to syllabic units can see their heads, but not, specifically, their dependents. The visibility of rhymal heads, as already mentioned, is evident from vowel harmony processes. The visibility of onset heads can be inferred from reduplication processes which reduplicated /pl/ as /p/, rather than as /l/ (cf. Steriade 1988).

The issue of complexity is relevant within syllabic constituents if relationships between segments are analyzed as α-α relations. It can easily be seen that heads have a greater array of choices than dependents in complex onsets. For starters, singleton onsets can generally be any consonant, while dependents (as is claimed in government phonology) can only be sonorants (cf. below). However, in complex onsets, we note that heads are also limited, in the sense that only obstruents can now be found in this position.[16] Nonetheless, heads allow greater complexity in that onset heads typically have an array of place and laryngeal options which are not available for sonorants in dependent positions.

In addition to making strong claims about the structural complexity of syllabic constituents, government phonology has also adopted strong restrictions on the kinds of segments that may occur in (especially) branching constituents. With reference to onsets, for example, (but this has been generalized to all head-dependency relations that holds between segments, including lateral ones), it has been claimed that dependents cannot be more complex than heads (see, especially, Harris 1994: 170–178); cf. (8). Since, in onsets, obstruents can combine with sonorants (as the only option allowed), it follows that, intrasegmentally, obstruents are more complex (in terms of their elemental make-up) than sonorants.

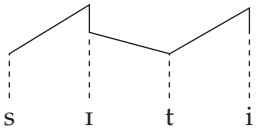
Government phonology shares with dependency phonology a crucial reliance on head-dependency relations. At the same time, Anderson and Ewen (1987) suggest that it may be the case that linguistic structures such as those encountered in phonology adhere to strict binarity, thus disfavoring or even disallowing multiple dependents on one head. Nonetheless, dependency phonology is not constituent-based. It appeals only to dependency relations as organizing relations. Corresponding to the representations in (17), we would have the following structures, although it must be added that the left-most representation, which exclusively consist of a node, are not obvious in the dependency model, which takes syllabic nodes to be strictly projections from segmental nodes:

(19) Possible syllabic constituents



A head–dependent approach does not exclude the possibility that one head could have more than one dependent (although this potential limitation is assumed in dependency phonology) nor, in fact, that one dependent has more than one head, although it is clear that the latter option, if allowed, leads to an important difference between dependency graphs and constituency graphs. Dependency phonology has proposed the use of improper bracketing for the representation of so-called *ambisyllabic consonants* (often suggested for single consonants following lax vowels in Germanic languages) (Chapter 6):

(20)



6 Empty-Headed Constituents and Lateral Head–Dependent Relations

Government phonology, as we have seen, allows the use of a constituent whose head position has no segmental content (so-called empty-headed constituents).

Let us first establish that empty-headed Os (or floating empty skeletal positions[17]) have been proposed in a variety of cases and different models where a vowel initial unit can behave as if it starts with a consonant (French *h-aspiré*; Turkish /ɣ/; Clements and Keyser 1983; Denwood 2006; Charette 2004), or in which such units trigger gemination of a preceding consonant (Lowenstamm 1999). Empty-headed rhymes are less familiar objects, although they also have been suggested in frameworks other than government phonology (Anderson 1982; Shaw 1993; Oostendorp 1995). The innovation of government phonology, however, was to introduce systematic constraints on the distribution of empty-headed rhymes.[18]

There are two partly independent reasons for introducing empty-headed rhymes in phonological representations, namely, to account for vowel–zero alternations and to account for phonotactic constraints.[19]

6.1 Vowel–Zero Alternations

Many languages display alternations between vowel and zero. A classical case is formed by the so-called *yer*-alternations in Slavic languages (see Scheer 2004 for extensive discussion). Of interest here are the typical conditions that cause a *yer* to be present or not (i.e. to delete). In general, a *yer* is present on the surface if it followed by a *yer* in the next syllable (Scheer 2004 offers extensive discussion of this alternation in Slavic languages). Hungarian also displays a vowel/zero alternation (Ritter 1995, 2006a) witnessed in *bokor* 'bush (nominative)' and *bokrok* 'bush (nominative plural)':

(21) a.

O	R	O	R	O	R	
x	x	x	x	x	x	
b	o	k	Ø	r	Ø	
[b	o	k	o	r]		'bush (nom. sg.)'

b.

O	R	O	R	O	R	O	R
x	x	x	x	x	x	x	x
b	o	k	Ø	r	Ø	k	Ø
[b	o	k		r	o	k]	'bush (nom. pl.)'

In this example we focus on the rhyme node that is underlined. We note that in (21a) this empty-headed rhyme is realized as a vowel if immediately followed by

another empty-headed rhyme. In (21b), however, the same unit is not realized and in this case we note it is followed by a contentful rhyme.

Kaye, Lowenstamm, and Vergnaud (1990) argue that the condition that accounts for the appearance of these vowels in Russian, Hungarian, and similar cases[20] can be generalized into a principle of grammar:

(22) An empty-headed rhyme is inaudible only if followed by a non-empty rhyme.[21]

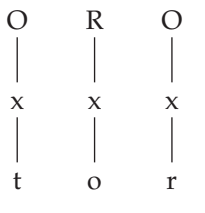
Kaye, Lowenstamm, and Vergnaud initially term this principle "Proper Government" and they show that the required relation can be blocked in certain cases, for example if intervening between the governee and governor we find an obstruent cluster. Proper Government is a head–dependent relationship, but it is non-structural, that is, it does not correspond to a constituent and the units involved in this relationship are not structural sisters. Following Scheer (2004), I will call such non-structural relations *lateral relations* and I will refer to them as *licensing relations*, ignoring various ways in which government phonologists have made a distinction between government (allowing emptiness) and licensing (allowing content); see Scheer (2004), Ritter (2006a) for discussion. I return to this point in Section 7.

The proposed treatment of vowel/zero alternations is compatible with various implementations. The fact that the vowels surface in ungoverned/unlicensed position can be the result of phonetic interpretation (perhaps to be seen as part of phonetic implementation), as in government phonology. Alternatively, the fact that they surface could be attributed to a repair rule that inserts a default element. Third, the element in question could be part of the lexical representation without being associated to the rhymal slot. Staying unassociated is on this analysis only possible if the slot is governed/licensed (cf. Scheer 2004 for a defense of this approach).[22]

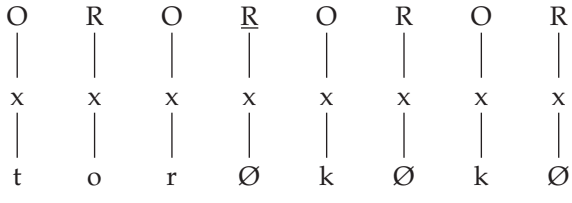
Irrespective of the precise details, all models using empty-headed rhymes to account for vowel/zero alternations share the claim that syllable structure is part and parcel of the lexical representation of morphemes and words rather than being derived on the basis of a linear sequence of segments by a syllabification procedure. This must be so, because otherwise the account proposed in this section simply does not work out, unless one were to adopt a special type of empty segment (a root node without content, perhaps) that would underlie the vowel/zero alternation. Postulating syllabic organization underlyingly raises the question as to whether syllabic organization by itself can be used distinctively, a possibility that is usually held to be unattested. However, if we consider the following pair in Hungarian, we have what seems to be a (near) minimal pair which is distinguished solely in terms of syllabic organization:

(23)	a.	torok	'throat (nom. sg.)'	torkok	'throat (nom. pl.)'
	b.	park	'park (nom. sg.)'	parkok	'park (nom. pl.)'

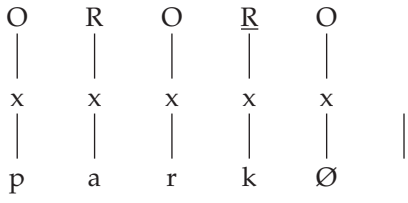
(24) a.



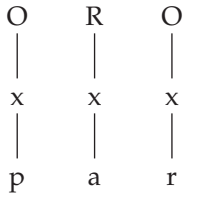
b.



(25) a.



b.



Examples of this kind illustrate that syllable structure can be distinctive.

6.2 Impossible or Improbable Phonotactics[23]

In the preceding sections some striking claims have been discussed by proponents of government phonology with regard to possible syllabic constituents. First, it has been argued that all syllabic constituents are maximally binary and, second, it has been stated that the segmental occupation of syllabic positions in branching onsets is limited: heads can only dominate obstruents and dependents can only dominate sonorants. Less controversial is the claim that in branching rhymes, heads must be vowels.

The question arises how such strong claims can be made if it seems so obvious that there are many languages in which onsets or rhymes exceed the number 2, or where complex onsets are of the type sonorant-obstruent (as in Slavic languages; cf. below). English allows tri-consonantal onsets (as in *spring* and *string*) as well as trisegmental rhymes as in (*spoon, warm*), taking a long vowel to be bisegmental, and also quadrisegmental rhymes, even in uninflected words (*paint, gold, world*).

Three responses are possible to the observation that ternary syllabic constituents are rare but not unattested. First, one might regard differences in frequency as irrelevant and claim that linguistic theories of synchronic states of languages must be able to represent, on an equal footing, whatever the history of languages produces. In this view, the claim that syllabic constituents are bounded is immediately falsified by an abundance of data. A second view could be that syllabic constituents of unbounded complexity must be represented on an equal footing with those that are bounded but not without having the grammar note that increasing complexity comes with an increasing number of violations of markedness constraints. This is the approach taken by proponents of optimality theory (Prince and Smolensky 1993; Kager 1999).

A third approach, taken by government phonologists, is that tendencies (in this case, apparent avoidance of unbounded syllabic constituents) suggest that there are *absolute grammatical constraints*. This does not entail that the falsifying data cannot be accounted for and will be ignored. Rather, it entails that these data require more abstract and more complex representations, which are available, although not without limits.[24]

An important discovery has been that phonotactic patterns that present apparent problems for government phonology are typically limited to word edges. It is well known, for example, that extra consonants can occur on the left or right periphery of words, leading to initial or final clusters that we do not encounter word-internally as syllable-initial or syllable-final clusters, respectively.

Extra consonantal options at the periphery of words are often referred to as *prependix*, *extraprosodic* (or *extrasyllabic*) *position*, and *appendix* (cf. Fudge 1969). Sometimes it is even proposed that *two* types of extra positions are allowed word-finally in order to account for so-called superheavy VXC syllables which are followed by a coronal appendix, as in Dutch *her-f-s-t* 'autumn', where the /f/ is the extra consonant producing the superheavy syllable (*herf*), while /st/ occupies the appendix position; cf. Trommelen (1983) and van der Hulst (1984).

In government phonology, the extra consonants are represented as what are called *degenerate syllables*, that is, syllables consisting of an overt onset and an empty-headed rhyme.[25] Similar ideas have emerged in moraic theories of syllable structure (Shaw 1993; Oostendorp 1995). In Kuryłowicz (1952), the peripheral degenerate syllables are stranded onsets, that is, not followed by an empty nucleus; this approach is also found in Polgárdi (1998) and, perhaps, Dell (1995).[26] In still other (usually more descriptively oriented) approaches, the extra consonants are not given any special structural status, and it is simply assumed that onsets and codas at word edges can be more complex. This approach receives a slightly more

theoretical status by developing the notion of extrasyllabicity (Rubach and Booij 1990a).

With the option of having extra word-peripheral consonants, a language can have CV as its *only* syllable type, allowing CCV word-initially and/or CVC word-finally, to be represented as C∅CV and CVC∅. In cases of this sort, initial CC clusters often contain sequences other than obstruent-sonorant, which is a further indication that these initial sequences do not form true onsets. In other cases, the permitted syllable is more complex (assuming that branching is a parametric option), allowing CC initially and VC finally. In such a case, word edges can also allow extra consonants, arriving at the possibility of word-initial triconsonantal clusters (always involving initial /s/) and word-final superheavy syllables (-VVC/-VCC), to be represented as C∅CCV[27] and CVXC∅.

The guiding idea of government phonology is that an elegant structural description of the phonological structure of words may call for postulating a somewhat abstract organization which contains units or terminal nodes that remain empty. In addition, government phonology, as shown above, appeals to principles (such as Proper Government) that control the distribution of these empty nodes.

I believe that the observation that certain complexities are limited to edges is quite crucial to counter the stipulative view that the syllable template for a language can be defined as the sum total of word-initial and word-final clusters separated by the set of vowels. I will now discuss a few cases which show that certain complexities are indeed confined to word edges.

An inspection of Dutch syllable structure (as found in Trommelen 1983; van der Hulst 1984) reveals that Dutch onsets can only exceed two in number when at the left word edge, in which case tri-consonantal clusters are allowed consisting of /s/ + obstruent + liquid:[28]

(26) stronk 'trunk'
splijt 'split'
sprong 'jump'

When such a tri-consonantal cluster is found word-internally (between two vowels), without the interference of a strong morpheme boundary, it is split up by a syllable division as follows:[29]

(27) mis.tral 'mistral'
es.planade 'esplanade'
Cas.tro 'Castro'

Independent evidence for this syllable division (as Trommelen 1983 points out) is that the vowel to the left of /s/ is *lax*, which is a signal that is being checked by a following tautosyllabic consonant (cf. van Oostendorp 1995). The claim that word-initial clusters need not be syllable-initial clusters can even be shown on the basis of seemingly well-behaved biconsonantal clusters consisting of an obstruent and a sonorant:

(28)	a.	gnoom	'gnome'	b.	Ag.nes	'Agnes'
		slaaf	'slave'		Os.lo	'Oslo'
		tjiktjak	'type of bird'		at.jar	'atjar'

The possible initial clusters /gn/, /sl/, /tj/ are split up when they occur intervocalically. This shows that the only real branching onsets are those consisting of an obstruent (excluding /s/) followed by a liquid (cf. Trommelen 1983).[30]

Another language that has word-initial clusters with more than two consonants is Polish (Rubach and Booij 1990; Gussman and Kaye 1993; Cyran and Gussman 1999; Rowicka 1999):

(29)	pstry	'mottled'
	bzdura	'nonsense'

Evidence for the word-internal syllabification of such clusters is harder to obtain given their scarcity in non-derived words (cf. below), but this is in itself an indication that the clusters may be restricted to the word-initial position. In addition to having such complex clusters, initial biconsonantal clusters appear to allow many combinations that violate the s*onority sequencing generalization* (Selkirk 1982; Clements 1990); but see Cyran and Gussmann (1999) and Scheer (2007) for restrictions that do exist:

(30)	ptak	'bird'
	scheda [sx]	'inheritance'
	skok	'jump'
	mnoz.yć	'multiply'
	lnu	'linen, gen.sg.'
	rtęć	'mercury'

Rubach and Booij (1990) note that the options for word-internal onsets are considerably restricted, suggesting that a cluster like [-rt-], allowed word-initially, is heretosyllabic word-internally: *kar-ty* 'cards'. This is very similar to what I reported for Dutch above.

The realization that clusters that exceed the size of two consonants, as well as clusters that also violate the sonority sequencing generalization, are restricted to word-initial position, frees the way to claiming that clusters that are grammatical at the left edge of words are not necessarily true onsets, which in turn triggers an investigation into their special nature.

Many phonologists have treated the extra options at the left edge of words by allowing an extrasyllabic consonant in that position (cf. Rubach and Booij 1990 for Polish), treated as stranded onsets by some (cf. Kuryłowicz 1952). This approach fails to impose any restrictions on how many consonants can thus be adjoined; there never seems to be more than one. Others have suggested that the apparent sequence of two consonants may in some case involve complex segments (s+C clusters; cf. van de Weijer 1996). The approach that is advanced in government

phonology (mentioned above) claims that the extra material involves degenerate syllables consisting of onsets followed by empty rhymes.[31]

With respect to rhymal structure we can make similar observations. Languages such as Dutch and English allow word-final rhymes that are rather complex:

(31)

Dutch		*English*
oogst	'harvest'	sixth
ernst	'seriousness'	blast(s)
vreemd	'strange'	clown(s)

In each case, we find lax/short vowels followed by up to four consonants (VCCCC), or tense/long vowels followed by up to three consonants (V_{tense}CCC or VVCCC). An inspection of word-internal syllables demonstrates to us that such very complex rhymes are rather rare when they are not word final. This leads us to the descriptive generalization that superheavy syllables (i.e. syllables ending in overcomplex rhymes) are limited to the right edge of words; this observation is also commonly made with reference to Semitic languages. The idea that word-final consonants may not belong to the core syllable was made explicit by Fudge (1969, 1987), who suggested a word-final constituent which he called the appendix, a notion that was subsequently adopted in other work (cf. Selkirk 1982, van der Hulst 1984).[32]

However, not all instances of impossible phonotactics regard edges. There are also instances of impossible phonotactics that involve intervocalic consonant sequences (i.e. interludes) that violate the 'syllable contact law' proposed in Murray and Vennemann (1983), which states that the edge consonant (i.e. onset head) should not be less sonorous than the preceding coda consonant. Vennemann and Murray state this as a dispreferred pattern, but government phonology has translated this into an absolute requirement. This, then, makes the heterosyllabic sequences in (32a) suspect:

(32) Turkish

a.	azmi	'resolution'
	kavmi	'tribe'
b.	metni	'text'
	kabri	'tomb'
	kudret	'power'

Given that it is generally assumed that Turkish does not allow complex onsets, the clusters in (32b) must be analyzed as heterosyllabic, which causes more violations of the syllable contact law. Denwood (2006) and Charette (2004) provide extensive analyses of Turkish which clearly demonstrate that all these alleged interludes behave as sequences of onsets (with intervening empty-headed rhymes).

A further complication is raised by syllabic consonants, which occur, for example, in English in unstressed syllables, but even in stressed syllables in many other languages; cf. Bell (1978):

(33) *English*

bottom	[batm̥]
bottle	[batl̥]
water	[wɑtr̥]

Here we seem to have the wrong segment type (here, a consonant) in the rhymal head position. Rather than relaxing what kind of segments can make up nuclei (at the phonotactic level), proponents of government phonology have suggested that such syllabic consonants are not rhymal heads, but, for example, a coda preceded by an empty-headed rhyme; see Scheer (2007).

Summarizing, we have considered the following circumstances:[33]

(34) a. An impossible onset
   i. Too many consonants (English, Georgian[34])
   ii. The wrong combination (Polish)
   b. An impossible rhyme
   i. Too many segments (English *spoon*)
   ii. The wrong segment (syllabic consonant, e.g. English *bottom*)
   c. An impossible interlude, i.e. violations of the syllable contact law (Turkish)

All these circumstances can be represented by appealing to empty-headed rhymes:

(35) ai. (sV)(prɪŋ)
aii. (lV)(nu)
bi. (tow)(nV)
bii. (bot)(tVm)
c. (a)(zV)(mi)

Can all cases of excess be analyzed in terms of just a single empty-headed rhyme that is licensed in terms of the principle in (22)? Things are not always so simple. Let us demonstrate this with onsets. If the principle in (22) is correct, this would predict that the complexity of initial clusters is limited to a sequence of two possible onsets. Cyran and Gussman (1999) explain that Polish might be more complicated than that, and they propose an additional principle (*Interonset Licensing*) which allows one extra onset (and a following empty rhyme). A similar case is made in Ritter (2006a) for Georgian, notorious for its onset complexity.[35] Final empty-headed rhymes can not be licensed in terms of principle (22). Thus, proponents of government phonology propose *final licensing* as an additional licensing mechanism.

Some languages even go further beyond all the complexities that have so far been discussed, most dramatically by allowing words to consist of sequences of consonants only, some of which are syllabic (i.e. syllable peaks) (Hoard 1978; Dell and Elmedlaoui 1985):[36]

(36) *Nuxalk* (formerly called Bella Coola; Bagemihl (1991)
xsc'c 'I'm now fat'
lxwtlcxw 'you spat on me'

Imdlawn Tashlhiyt Berber (Dell and Elmedlaoui 1985, 2002)
tftkt 'you suffered a sprain'
ssrksxt 'I hid him'

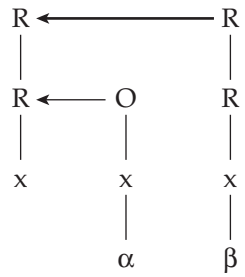
With respect to the Nuxalk facts, it has been suggested that such sequences should be taken as evidence for the claim that syllables can consist of just onsets (Hockett 1955), that the segmental string lacks syllable structure (Newman 1947), or that there is only partial syllable structure (Bagemihl 1991). Dell and Elmedlaoui maintain that words in Imdlawn Tashlhiyt Berber are completely syllabified, an analysis which requires that all consonant types can function as syllabic peaks.

7 Licensing of Marked Constituents

In the preceding sections we have seen several non-structural (that is, lateral) head-dependent relations, referred to either as government or licensing relations. We have considered "Proper Government" (22) which licenses empty rhymes, but other licensing mechanisms have been proposed (Kaye (1990a), Charette (1990), and later works).[37] Taken together, these mechanisms suggest a generalization that is explored in van der Hulst (2006b).[38] I will summarize the proposal here. The central idea is that a marked constituent (branching or empty-headed) cannot occur freely but must be licensed by constituents that follow, which must be non-empty. Specifically, it would seem that all the licensing/government principles/parameters that government phonology has been proposing can be subsumed under a single generalization: *marked* constituents must be locally licensed by following *contentful* constituents (only R in case of marked O, and both O and R in case of marked R).

An example of the general licensing scheme is as follows:

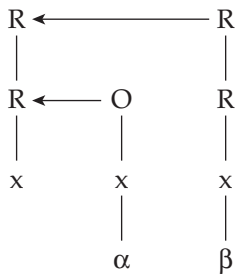
(37)



This scheme says that an empty-headed rhyme must be followed by a contentful onset and a contentful rhyme. The second requirement was stated explicitly in government phonology under the heading of "proper governed." The original idea was that an empty-headed rhyme that was not properly government would have to become audible as a matter of phonetic interpretation. This view accounts

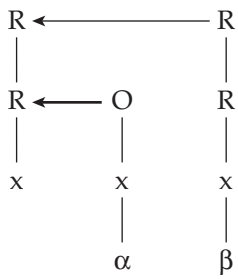
for what others might see as a repair rule (insertion of an element to make the rhyme audible):

(38)



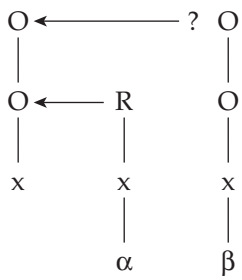
The second requirement is noted in Charette's (1990) treatment of schwa-deletion in French. Words like *dehors* [dəɔr] 'outside' do not allow the schwa to become silent. The idea in this analysis is that a French schwa is represented as an empty rhyme which is not governed, or, as I would say, licensed. In this type of example, we see that the schwa must be audible in hiatus even when there is a following non-empty rhyme and the reason seems to be that there is no onset, or an empty onset:

(39)



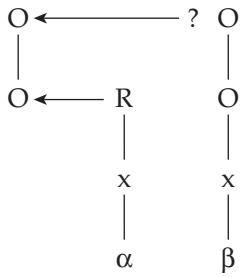
Interestingly, it is not obvious whether empty onsets also require double licensing. Here I need to distinguish between an empty-headed onset (with a skeletal point, such as h-aspiré in French) and onsets that are truly absent, that is, lacking a skeletal point. It is probably trivially true that no cases of two consecutive empty-headed onsets have been reported, but that does not mean that such a sequence is illegal, given that such units appear to be fairly rare in the first place. Hence, the case for the OO relation in (40) is weak:

(40)



What about the absence of O, or at least the absence of an O that has a skeletal position?

(41)



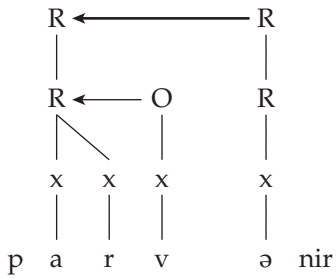
The OR relation appears to be supported by the fact that empty-headed rhymes seem to always require the presence of a contentful onset, which boils down to the non-existence of an entirely empty syllable (i.e. OR sequence). It is here that we see the advantage of adopting "pointless" Os because unless we assume these units the constraint against entirely empty syllables could not be stated as part of the licensing paradigm that this section develops.

The OO relation in (41) would be warranted if it were true that languages avoid double hiatus, for example, a string like /. . . a – o – i . . ./ (which does occur in the English word *maoist*). A constraint against double hiatus has not, to my knowledge, been proposed.[39]

Having discussed licensing constraints on empty-headed constituents, which, as shown, are well supported for rhymes and less so for onsets (whether with a skeletal point or without it), I now turn to the licensing requirements on branching constituents.

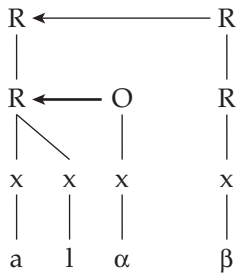
Let us first consider branching rhymes that end in a coda consonant. Van der Hulst (2006b) suggests that these two must be followed by a contentful rhyme, which can be demonstrated by considering certain facts of French. Charette (1990) argues that the schwa in (some varieties of) French cannot be silent in words like *parvenir*. She proposes an account which is different from the one I suggest here.[40] The schwa is required to be present so that it can license the /v/ to govern the preceding coda /r/. In the spirit of Harris (1997) I suggest that the licensing goes directly from the rhyme with the schwa to the preceding branching rhyme.[41]

(42)



Let us now ask whether a branching rhyme must also be licensed by a following contentful onset.

(43)



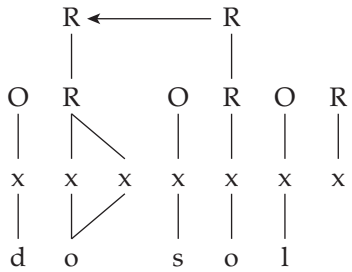
It would seem that (43) is well motivated and is, in fact, known as "coda licensing" in government phonology (Kaye 1990a), which instead construes the relevant mechanism as the required presence of an onset following a coda consonant.

A second case of branching rhymes might be rhymes with a long vowel. I will assume here without discussion that all long vowels are necessarily birhymal. If we were to reject that idea and allow monorhymal long vowels, it would seem at first sight that we could explain the well-known closed syllable shortening effect in terms of the licensing requirements on branching rhymes that we saw for closed syllables. Closed syllable shortening is witnessed by alternations in, for example, Yawelmani (Kaye 1990b; Yoshida 1993; van der Hulst 2006b):

(44) do:s - ol 'report (dubitative)' dos - hin 'report (non-future)'

Kaye (1990a) proposes *not* to account for the appearance of a short /o/ in *doshin* by appealing to a shortening that is caused by the vowel appearing in a closed syllable and a constraint on rhymes not exceeding two positions. His reluctance to accept this traditional account stems from the fact that in government phonology, syllabic affiliations of segments must remain unchanged. In addition, there is no compelling argument for saying that the vowels that shorten are in closed syllables; shortening takes place before any two consonants, no matter what their sonority slope. In other words, resyllabification is not allowed in government phonology. The shortening effect is instead attributed to the fact that a branching rhyme must be licensed by a following contentful rhyme. This is the case in *do:sol*:

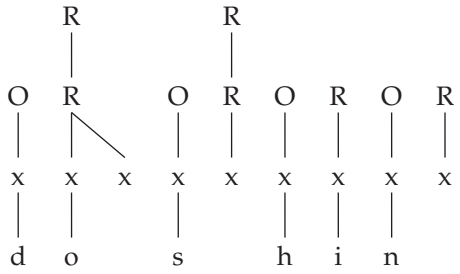
(45) a.



/do:s/ - ol/ -> do:sol

However, in the case of underlying *do:shin* the underlying long vowel is followed by an empty-headed rhyme, and this would imply that the branching rhyme that supports the long vowel is not licensed:

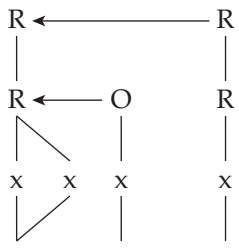
(46)



/do:s - hin/ -> doshin

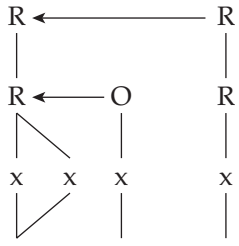
Analogous to (42), it would seem, then, that long vowels *as branching rhymes* must also be licensed by a following overt rhyme:

(47)



Let us now ask whether a long vowel as branching rhyme must also be licensed by a following contentful onset (analogous to 43):

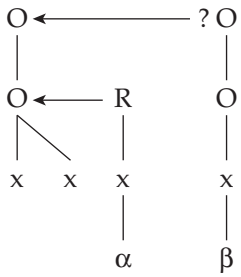
(48)



(48) says that if long vowels are indeed branching rhymes, we would expect that they must be followed by an overt onset, but there is no support for that restriction. Hiatus after long vowels is quite common. In English, vowels to the left of hiatus are, in fact, obligatorily long. The distribution of long vowels, it would seem, suggests that long vowels are *not* branching rhymes.[42]

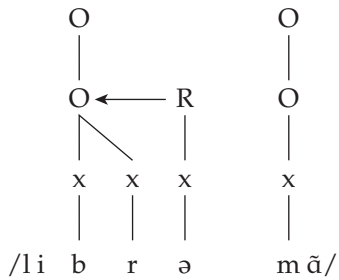
Finally, to complete "the licensing paradigm" we will address the need to license branching onsets; recall that we found little support for the notion that empty onsets require licensing by a following overt onset (cf. 41, lower arrow), while they do require licensing in terms of a following overt rhyme (cf. 41, upper arrow):

(49)



Again there seems to be no evidence for the OO relationship. Strings like /bri.o/ do not seem problematic in languages that allow both branching onsets and empty onsets. However the OR relationship is well motivated, as demonstrated by the other case that Charette (1990) discusses as an instance of "license to govern," the idea being that /b/ in (50) must be licensed to govern its dependent /r/ by an audible nucleus. Hence, after a branching onset schwa, must be pronounced. In line with the system of licensing developed here, I interpret this case as direct licensing of a branching constituent.

(50)



The following table summarizes the results:

(51)

	Homogeneous licensing	*Heterogeneous licensing*
Empty Rhyme	RR (proper government)	RO (*dehors*-case)[43]
Empty Onset	OO [not needed]	OR ("*empty syllable")
Branching Rhyme	RR (*parvenir*-case)	RO (coda licensing)
Branching Onset	OO [not needed]	OR (*librement*-case)

Van der Hulst (2006b) suggests that the apparent non-existence of OO licensing would find a principled explanation if one assumes that there is no O-projection level, comparable to the R-projection level. Only the latter can be independently motivated by the fact that R is the head of the OR package.

In conclusion, if we assume that empty and branching syllabic constituents are *deviations* from the unmarked contentful, non-branching case, it would seem that

all the licensing/government principles/parameters that government phonology has been proposing can be subsumed under a single generalization: *marked* constituents must be locally licensed by following *contentful* constituents (only R in case of marked O, and both O and R in case of marked R).[44] This makes intuitive sense because otherwise proliferations of marked, that is empty-headed and branching, constituents, would lead to excessive consonant sequences.

We have seen two motivations for the use of empty-headed rhymes: vowel/zero alternations and impossible phonotactics. While in the former type of analysis, empty-headed rhymes are motivated in terms of observable alternations, this is not so in the latter case. For this reason alone, one might call their use for these types of cases unacceptable. There is nothing inherent in a dependency approach, not even one that limits itself to binarity, which compels the postulation of empty-headed rhymes in case phoneme sequences overstep the boundaries of what is considered to be unmarked.

8 Other Recent Developments

Neither dependency phonology nor government phonology are static models. Since their inception modifications have naturally been proposed. I consider radical CV phonology primarily a development (van der Hulst 2005a) of dependency phonology, but I also refer to ongoing work by Anderson (2004, 2006), and somewhat more distantly related work by Smith (2000), Humbert (1995), Botma (2004) and van der Torre (2003).

Radical CV phonology pursues the idea that the internal and external syntax of phonological segments can be represented in terms of just two primitives (labeled C and V) which are intrasegmentally grouped into three sets (place, manner, and laryngeal) and which extrasegmentally represent the syllabic constituents into which segments are grouped. Although monovalent in nature, the two primes are clearly antagonistic or polar.[45] It is possible to conceive of the elements C and V as subprimal units which, in conjunction with the group labels (place, manner, laryngeal) *define* the six elements that might be sufficient to characterize all possible phonological contrasts.

There have also been developments that have led to interesting varieties of government phonology. Van der Hulst and Ritter (1998, 1999b) offer a modification of some aspects of government phonology (such as removing the distinction between nucleus and rhyme). In addition they develop a typology of the various licensing relationships that can be distinguished. Scheer (2004) discusses both government phonology and varieties in some detail.

Second, I will mention an influential idea put forward in Lowenstamm (1996; developed in 1999, 2003) which is the proposal to universally rule out branching syllabic constituents. This effectively reduces all languages to strict CV languages. In this proposal, all apparent codas are onsets, and all long vowels are bi-rhymal. Moreover, all branching onsets are sequences of onsets with intervening empty headed rhymes. Lowenstamm (2003) proposes to treat complex onsets differently,

namely as complex segments which are represented under a single skeletal position.[46] Lowenstamm's strict CV idea has been adopted in Scheer (2004), who offers a very detailed application of this approach.[47] Scheer refers to his version of this approach as a "lateral approach" to phonology because rather than appealing to (syllabic) constituency and hierarchy, all relationships involve lateral connections between O and R nodes (which he notates as C and V, the distinction between syllabic node and skeleton no longer being necessary).[48] I also refer to Polgárdi (1998), Rowicka (1999), and Szigetvari (1999, 2000) for applications of this approach.

Third, Pöchtrager (2006) offers a rather different version of government phonology in which the role of former elements (in particular, the elements H and ʔ, responsible, among others, for the difference between fricatives and stops[49]) are taken over by more elaborate syllabic configurations. A different approach that also eliminates the elements for "continuancy" is offered in Ritter (1997). Here, continuancy is expressed in terms of intrasegment headedness (stops are headed and fricatives are headless), which has the advantage that it simply extends the use of the contrast between headedness and headlessness from vowels to consonants. In vowels, others had already argued that the difference in "ATR" could be expressed in this way (cf. Harris and Lindsey 1995).

The results obtained in these alternative approaches, as well as the one presented in this chapter, are not entirely incompatible. In some respects they are, but in others it is likely that all this work contributes to a dependency-based, or indeed head-driven theory of phonology.

9 Concluding Remarks

The present chapter has offered an overview of an approach to phonological structure that relies heavily on the role of head-dependency relations. In addition, it uses only monovalent primes and also, in some varieties, "element grouping." It is interesting to note that all three theoretical devices (headedness, monovalency, and grouping) were originally introduced in the early 1970s by proponents of dependency phonology, only to emerge in the mid-1970s and 1980s as independent developments within mainstream generative phonology.

Varieties of dependency-based phonology (dependency phonology, government phonology, radical CV phonology, and so on) are primarily theories about phonological representations, which are, as such, compatible with various views on the *derivational* aspect of phonology, that is, dealing with invariance, namely, alternations (both allophonic and allomorphic). Dependency phonology has essentially adopted a rule-based approach, in principle allowing for extrinsic ordering, but in practice having been insufficiently applied to the kind of data that seem to require this mechanism.[50] Government phonology has been advocated as a "no rule" approach, seeking an account of phonological structure and invariance that is entirely constraint-based. However, as pointed out in Section 2, the mechanism of "phonetic interpretation" seems to function as a repair component, extending the domain of elements and assigning a null interpretation to elements that are

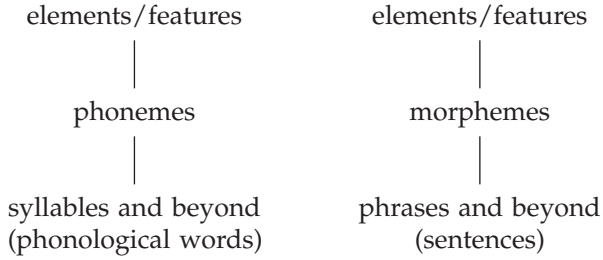
not licensed. Van der Hulst (2011) points out that we can also implement government phonology (or other variants of the models discussed here) in a declarative framework which accounts for invariance by unifying lexical representations (which themselves are very specific constraints) with constraints that "add" information, or in terms of lexical representations that contain disjunctions (called "hyperspecification," underspecification being a specific case); cf. Scobbie, Coleman, and Bird (1996).

Yet another way of dealing with invariance is to invoke an optimality theory style approach. A combination of, for example, government phonology and optimality theory can be found in Polgárdi (1998) and Rowicka (1999).[51] What these works clearly demonstrate is that issues of representation are, in principle, independent of issues of derivation, although it seems obvious that no derivational theory can even be conceived in the absence of a solid and explicit representational theory. It seems to me that the models discussed in this chapter primarily seek to develop such a theory. To use an apt phrase of Brandão De Carvalho (2002) "constraint-based theories need theory-based constraints."

An aspect of optimality theory that has not been explored in this chapter is that ranking between any two constraints can be interpreted as an instance of dependency relations between these constraints. It would, therefore, be in the spirit of dependency-based approaches to explore the use of dependency relations between constraints.

Crucial reliance on the notion "head" (and "dependent") suggests an analogy between phonology and syntax. I believe that the analogies between syntax (morphotactics) and phonology (phonotactics) are real and deep (van der Hulst 2005b, 2006c; contrary to Bromberger and Halle 1989).[52] Both systems form a complementary and essential part of the grammar of human languages and it would seem that both are relying on the same logic, that is, a system of intermediate primitives (morphemes and phonemes, respectively) and a combinatorial system of a specific sort (binary-bounded, headed). The intermediate primitives can be analyzed into a system of features or elements, which are, then, the ultimate primitives, and another combinatorial system:

(52)



With good reason, linguists have argued that both combinatorial systems make reference to the head–dependent asymmetry (whether taken as basic or augmented to a constituency) and perhaps also share binarity. The idea that different parts of the grammar rely on the same kind of computational mechanisms has been

termed "structural analogy" by John Anderson (e.g. Anderson 1987, 1992, 2006; van der Hulst 2005b). We still have different modules (i.e. phonotactics and morphotactics), but rather than expecting that these modules are organized totally differently (just because they are autonomous under some interpretation), it makes more sense, Anderson suggests, to adopt the working hypothesis that the human cognitive system replicates simple and successful procedures in different modules.[53]

If this view is correct, we require a notion of head (and dependent) which truly generalizes over both domains and not one that has been tailored to morphotactics (only to then claim that it does not apply to phonotactics). The claim that the combinatorial system for morphotactics is recursive, while that for phonotactics is not, is irrelevant to the analogy. Phonotactics primarily caters to phonetics (which has a rhythmic, iterative structure), while morphotactics primarily caters to semantics, which is inherently recursive. The need for embedding (i.e. recursion) originates in semantics, phonetics requires no more than iteration (cf. van der Hulst 2010). This does not make morphotactics the core of human language. It means that recursion is an available cognitive device, used where needed.

ACKNOWLEDGMENTS

I would like to thank John Anderson, Gienek Cyran, Irina Monich, Nancy Ritter, John Rennison, Tobias Scheer, Peter Szigetvari, Norval Smith as well as the editors of this handbook for very helpful comments on a prefinal draft of this chapter. Needless to say that they do not necessarily agree with all that I say here.

NOTES

1 I discuss some important varieties briefly in Section 8. See van der Hulst (2006a) and Ritter (2006) for overviews of dependency phonology and government phonology, respectively. Other introductions to government phonology can be found in Harris (1994), Brockhaus (1995), Ritter (1995), Scheer (2004), Cyran (2010) and to DP in Ewen (1980), Anderson and Ewen (1980a), Lass (1984), Anderson and Durand (1986, 1987b), Anderson, Ewen and Staun (1985), Ewen (1986), Durand (1990), Ewen (1993), and Anderson (2004). For both government phonology and dependency phonology see Ball (1997) .

2 Levin (1985) indeed proposed a theory of syllable structure which suggests a structural analogy between syllables and syntactic phrases as headed constituents. The parallelism between syllables and syntactic phrases is, however, an older observation; cf. Pittman (1948), Pike (1943b).

3 By analogy, in syllable structure, the nucleus is atomic (it contains a vowel), while onsets and codas are potentially complex. It might be countered that the nucleus can contain a long vowel or diphthong and if we wish to maintain that nuclei are atomic this would mean that such entities are complex intrasegmentally, or that they form two nuclei.

4 In dependency phonology and government phonology primes are single-valued (monovalent, unary) entities, referred to a *components* and *elements*, respectively. Here I use the latter term.
5 The need for both root nodes and skeletal position is debatable. cf. Selkirk (1990) and Section 5.
6 In this chapter, I do not discuss constraints of the type c, d, and e, although specifically the latter (the sequential content constraints) too can be construed as non-structural, lateral head–dependent relationships, whereas reference to head or dependent status places no role in c and d.
7 As is well-known, this need depends on the abstractness of underlying forms. In general, the practice in dependency phonology has been to not embrace the excessive abstractness of standard generative phonology.
8 Another early proponent of simplex features is Sanders (1972b).
9 It is interesting that dependency phonology (in a specific development called "radical CV phonology"; van der Hulst 2005a) and developments in government phonology (e.g. Ritter 1997) have converged on the hypothesis that no more than six monovalent primes are necessary to account for all contrasts and "processes" in all the world's languages. For general motivations for unary elements see Anderson and Ewen (1987), van der Hulst (1989), Harris and Lindsey (1995).
10 An additional argument is provided in Section 4.
11 The idea that syllabic constituents can branch is abandoned in the so-called "strict CV" version of government phonology (Lowenstamm 1996, 1999; Scheer 2004); see Section 8.
12 Below, we will consider whether such empty-headed constituents can be branching, that is, have a contentful dependent. Also note that government phonology assumes a "nucleus node" which is dominated by the rhyme node. I dismiss that option and discuss the issue below. Finally, being constituent-based government phonology uses skeletal points which, put simply, are the terminal nodes of the syllabic constituents.
13 Not allowing a distinction between a pointless and a pointed empty constituent would allow the use of bare O and bare R for empty-headed, that is, contentless constituents, but nothing, except the saving of ink, depends on that difference.
14 A string of any length can be represented in terms of an *unbounded* binary branching structure. Hence the strict claim made by government phonology should be stated such that it allows only *bounded* (in the sense of non-recursive) constituents. Kaye (1990a) states this by requiring that a head c-commands all its dependents.
15 Initially, dependency phonologists designated the sonorant consonant as the head of a branching onset, claiming that in both onsets and rhyme maximal sonority determines headedness. In Anderson (1986) it is suggested that obstruents are heads of onsets so that now the generalization is that in both syllabic constituents the most typical or optimal and most preferred segment type is the head. This view, which was also adopted in government phonology (but not in Scheer 2004), is more appropriate in that, in general, the head determines the "nature" of the constituent of which it is the head. This being said, it must be allowed that in non-branching onsets or rhymes non-optimal segments (such as sonorant consonants) can be heads as well; I return to this issue below.
16 This kind of "polarization" can be seen as a specific instantiation of the head–dependent asymmetry principle in that in head–dependent combinations, heads "retreat" to the optimal choices, to enhance the contrast with their dependents. We also see this in

stress feet when head syllables bar schwa-vowels, which are precisely the vowels that occur in dependent syllables (as in English).

17 If we regard skeletal positions as terminal nodes of syllabic constituents, then, these positions cannot exist in the absence of syllabic constituent structure. In other models, skeletal positions have been granted a more autonomous status (implicitly or explicitly) on the assumption that the syllabic grouping of skeletal positions is not part of the lexical representation, but done later due to a syllabification process. The model discussed here follows government phonology in explicitly adopting the view that syllabic grouping is inherent in the lexical representation. Arguments for this position will be provided below.

18 It has been suggested that the need for such constraints is analogous to the need for constraints that limit the distributions of entities in syntax such as traces and silent pronominal elements. In both cases we are dealing with mechanisms that control the distribution of empty categories.

19 For a discussion of how empty-headed rhymes interact with stress assignment, see Szigetvári and Scheer (2005).

20 See Scheer (2004) for several other cases and relevant references.

21 This principle would disallow any final empty rhyme; these do appear in some government phonology analyses and thus require an additional licensing mechanism. I return to this issue in Section 6.

22 In Russian there are two different *yer* vowels which means that we cannot represent both in terms of an empty-headed rhyme since this would make them phonologically identical. Without undermining the principle in (22), we would have to differentiate both *yer* rhymes by assuming differentiating elements that are not associated. Scheer (2004) generalizes this approach to all cases of empty rhymes, which has the advantage that it is easier to deal with the fact that empty rhymes do not get realized the same way in all languages. A different formalization of this same idea would be to use the "old" idea of disjunctive representations (cf. Hudson 1974) which allows a representation of units that are present only under specific conditions.

23 This section is a modified version of a section in van der Hulst and Ritter (1999b).

24 A sharp difference between dependency phonology and government phonology has always been that proponents of the former, while making available the essential representational apparatus that both models use (binary, headed constituents) have not been inclined to translate tendencies into absolute constraints. Thus, dependency phonology has been focused more on introducing fundamental concepts and less on imposing restrictions. This does not exclude the possibility that such restrictions could be added, or built in (cf. van der Hulst 2005a for a discussion of this point).

25 Rennison and Neubarth (2003) have a proposal involving a notion "strength" which determines which consonant can precede or follow which other consonant.

26 Cf. Dell (1995: 19): "we are assuming in effect that words such as *garde, marbre*, and so on, end in a degenerate syllable, that is, a syllable whose rime consists of a nucleus which is not associated with any distinctive features." A few lines later, however, Dell says that "the final empty nucleus does not belong to the lexical representations."

27 Kaye (1992) proposes to represent clusters that start with /s/ as ∅/s/C(C)V which explains, for example, why in Spanish an initial /ɛ/ vowel appears in such cases.

28 See Ewen (1982) and van de Weijer (1996) for analyses in terms of complex segments.

29 There are problematic cases such as the word "extra" [ɛkstra] which either has too much in its coda or too much in its onset.

30 For a government-based analysis of Dutch syllable structure cf. van der Hulst (2003).
31 This recasting of extrasyllabicity in terms of syllables with an empty-headed rhyme might help to explain why extrasyllabic consonants are more limited on the left-hand edge of words. See, in this connection, Lowenstamm (1999).
32 Morphological structure is relevant in these cases. Affixes can create additional complexities that require recognition of the fact that these affixes do not belong to the same phonotactic domain as the stem, but instead form an independent domain.
33 In all cases discussed, the generalization that certain complexities only occur at word edges must be qualified by saying that the notion "word" here refers to non-compounded, non-prefixed words, and words that do not contain certain classes of suffixes, such as so-called "level II" affixes and inflectional affixes. Exactly how to characterize the scope of the syllabic domain is not a trivial matter; however; see Kaye (1995).
34 See Butskhrikidze (2002) and, for a government phonology analysis, Ritter (2006a).
35 Interonset Licensing creates a head-dependency relationship between two onsets, the first one (the head) being more sonorous than the second one. This creates a "pseudo-onset." It is assumed that an empty headed rhyme caught within such a relationship is allowed to remain silent without being properly governed, that is, licensed by a following contentful rhyme (cf. 22). This then allows [O∅O∅OR . . . , where ∅ stands for an empty-headed rhyme.
36 In both cases we need to be careful in that transcriptions may not reveal epenthetic vowels. Dell and Elmedlaoui (1985, 2002) offer detailed discussion of the status of such vowels.
37 Collections of government phonology work can be found in two special issues of the journals *Phonology* (1990, Issue 7 and *The Linguistic Review* 2006, Issue 23/4. See also many articles in Ploch (2003). A number of dissertations have been produced coming from SOAS (London); also see the *SOAS Working Papers in Linguistics* and Cyran (1998). Collections of DP work are: Anderson and Ewen (1980b), Durand (1986), Anderson and Durand (1987b). There are, in addition, many other book-length government phonology studies of single languages (often originally dissertations) such as, for example, Yoshida (1990), Charette (1991), Brockhaus (1995), Ritter (1995), Rhee (2002), Polgárdi (1998), Rowicka (1999), Dienes (2000) and Szigetvari (2000).
38 This proposal does not cover cases for which government phonology invokes final licensing.
39 In fact, since long vowels are represented as birhymal in this approach, any long vowel that is not preceded by an overt onset (as in English *eel*) presents a case of two consecutive pointless Os.
40 Also see Charette (1992, 2003).
41 Charette (1990: 240) claims that schwa deletion is not only impossible if an "impossible onset cluster" like "vn" would arise, but also in cases like *tourterelle* 'turtle dove' where the resulting combination /tr/ would be a fine onset. I am aware of the fact that the facts of schwa deletion are not always clear and it is therefore necessary to support the RR relation in (39) with additional cases.
42 I refer to Lowenstamm (1996) for accounts of closed syllable shortening effects under bi-rhymal analyses of long vowels. Lowenstamm suggests that an unlicensed empty headed rhyme is only made audible if preceded by an overt onset. This does not apply to the second (empty) rhyme of a long vowel which is not preceded by an overt onset.
43 Van der Hulst (2006b) provides further motivation for this licensing case with reference to what is called "resolution" or "reduction" in the government phonology literature.

If a stem ends in a empty headed rhyme and a following suffix starts with an empty onset, a violation of heterogeneous licensing of empty rhymes, both empty constituents are removed.

44 An issue that needs more research is the question whether licensors, in addition to being non-empty, are also preferably non-branching.

45 One might summarize the RCVP view on primes as follows: all primes are monovalent; there are only two (polar) primes. This sounds as the ultimate compromise between unary and binary approaches.

46 This proposal is similar to Duanmu's CVX theory (Duanmu 2008) who, unlike Lowenstamm allows branching rhymes (VX). Duanmu also argues that alleged complex onsets are complex segments.

47 This work also contains a very useful and detailed discussion of other varieties of government phonology. See Cyran (2006) for an extensive review.

48 This approach is, in certain respects, prefigured in Takahashi (1993, 2004), who extends Harris' (1997) notion of licensing inheritance.

49 Jensen (1994) had already questioned the need for the element ʔ.

50 This does reflect a bias against overly abstract underlying representations. I would say that DP favors a fairly concrete phonology and thus will tend not to rely on extrinsic orderings.

51 Independent of OT, the notion of constraint ranking was invoked in Charette (1990) and Cyran (1996), both government phonology accounts.

52 See the articles in R. Bermúdez-Otero and P. Honeybone (2006).

53 There is, in my opinion, no need to believe that these mechanisms are linguistic rather than cognitive (cf. Anderson 2006).

17 The Acquisition of Phonology

KATHERINE DEMUTH

1 Introduction

The field of phonological acquisition has changed and developed over the years, in tandem with various developments in phonological theory. Like the field of phonology itself, the study of phonological acquisition began with an examination of segments and words (Yeni-Komshian *et al.* 1980). Much of this work found a large amount of individual variation in children's early productions. This led to a more cognitive approach to phonological acquisition, where issues of individual strategies predominated, with little in the way of predictions about the course of phonological development. To be sure, the field was also small, and there was a limited amount of data available for investigating issues of phonological development. Over the past 15 years this situation has begun to change, and there has emerged a growing amount of literature examining the acquisition of phonological structure at the higher levels of the syllable and the prosodic word. This new research focus has been stimulated in part by developments in prosodic phonology and phonological theory that provide a framework for investigating children's early language productions in terms of a constraint-satisfaction problem, rather than as a rule-based system. This has coincided with the recent availability of new longitudinal, phonetically transcribed, computerized corpora from children learning many different languages, ideal for addressing issues of early phonological development. At the same time there has been increasing contact with other fields, including infant-speech perception, speech and hearing sciences, and computational modeling. The purpose of this chapter is to review these developments in more detail, providing the reader with a sense

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

of where the field began, where it is today, and some of the promising new directions for the future.

2 Early Studies of Phonological Development (1870s–1960s)

2.1 From Descriptivist to Behaviorist Approaches to Phonological Acquisition

Scientists and parents alike have long been interested in language development (e.g. Darwin 1877; Deville 1891). Many of these early works constituted diary studies, where the observant parent wrote down the child's words or utterances. One of the most influential in terms of phonological acquisition was Leopold's (1939–1949) study of his daughter Hildegard's language development from birth till 2 years. The fact that she was developing bilingually (English and German) made this study all the more interesting. Although there are obvious limitations to diary study methods, such as the objectivity, accuracy, and training of the observer, they have provided some of the most complete records of individual children's early phonological development. These studies typically include lists of children's early words, giving some indication of early word and syllable shapes, the syllables and segments that were omitted, and the types of segments that were modified. Such diary data are therefore ideal for investigating segmental and word-level phonological development. They are less amenable to examining the use of words in larger prosodic context, or for examining the development of morpho-phonological interactions.

The individual case study nature of these early data sources, and the lack of any overarching theoretical perspectives at the time, led to a descriptive view of the data. This changed with the rise of behaviorist perspectives after World War I, with a shift in focus to more large-scale controlled studies of language development. The most influential normative study of this period was Templin (1957), where 430 children aged 3–8 were assessed for their production and discrimination of segments, as well as receptive vocabulary and sentence production skills, laying the groundwork for future research on normative language development.

With the rise of behaviorism, Leonard Bloomfield (1933) began to see children's language development as being determined and shaped by the environment in terms of imitating what was heard in order to communicate one's thoughts and desires. He proposed that the child acquired words separately in comprehension and production, and only later brought the two together (cf. Kiparsky and Menn 1977 and Smolensky 1996b for later views). In keeping with behaviorist thoughts at the time, Skinner (1957) also had the notion that the child's ill-formed utterances would be corrected through explicit negative evidence, something that further research has shown to be rare and ineffective at modifying children's phonological or morphological representations.

2.2 *From Structuralism to Generative Approaches to Phonological Acquisition*

In the meantime, the groundwork for formal linguistics was being laid in Europe. Ferdinand de Saussure (1916) was the first to move beyond surface segmental contrasts to propose a more abstract unit, the phoneme. This was incorporated into the study of Slavic languages by the Prague Circle in the 1920s and 1930s, where Nikolai Trubetzkoy (1939) and Roman Jakobson (1941) went beyond the notion of phonemic inventories to think of language in terms of phonological systems. Roman Jakobson (1941) proposed that all children begin the process of language acquisition by producing a maximally different set of unmarked contrasts (e.g. /p/, /n/, /ɑ/). This specific proposal has never been verified at the segmental level (e.g. Velten 1943). However, the notion that phonologically unmarked structures (e.g. stop consonants, CV syllable structures) will be the first to be acquired, has generally been upheld (cf. Lindblom 1992; Locke 1983), and has had a recent resurgence of influence in the field (see Section 5.1).

Noam Chomsky's (1957) publication of *Syntactic Structures* began to develop the notion of abstract (underlying) representations in syntax. This was followed by *Aspects of the Theory of Syntax* (1965), where Chomsky argued that language was too hierarchical and noisy to be learned merely from interactions with the world. Rather, he proposed that learners were constrained by a set of universal underlying principles that guided the language learning process. This nativist (or rationalist) approach to language was in direct contrast to the previously held behaviorist approaches to language acquisition.

The structuralist focus on surface phonological feature contrasts had been an advance over previous surface segmental approaches to phonology, but it had little to say about suprasegmental phenomena such as stress. This led to the development of generative phonology in the mid 1960s, with *The Sound Pattern of English* (Chomsky and Halle 1968) providing the foundation for early generative phonological theory. As in case of syntax, it was realized that an adequate theory of phonology must characterize what a speaker (and learner) knows about the sound system of the language at an abstract level of phonological representation. This laid the groundwork for both more formal linguistic, as well as more cognitive/biological approaches to phonological acquisition.

3 The 1970s and 1980s: Generative vs. Functionalist Approaches to Phonological Acquisition

The 1970s saw the rapid rise of studies in the acquisition of phonology. Some took a generative approach, formulating a comprehensive set of phonological rules to account for children's early productions (e.g. Smith 1973). Others took a more data driven approach, explaining individual difference in terms of cognitive and biological self-organizing systems (Ferguson and Farwell 1975; Lindblom 1992;

see Vihman 1996 for a review). This section highlights the important contributions and limitations of both.

3.1 Phonological Rules (and Constraints)

The first comprehensive longitudinal case study of a child's phonological development came from Neil Smith's (1973) diary study of his son Amahl. This was a landmark study in terms of providing a generative, rule-based account of the child's phonological development between 2;2 and 4 years. Smith proposed that the child's productions were not a result of misperception, but rather the child's own rule-based system. Further support for this proposal was the finding of widespread across-the-board phonological processes, similar to some of those found in both synchronic and diachronic phonology. One of the interesting phonological findings was that of chain shifts (*puzzle* > *puddle* > *pickle*), analyzed further by Macken (1980). This study has had a lasting effect on the field by publishing all the data, providing a rich source of information for reanalysis as phonological theory has developed over time. A few other studies also took a comprehensive rule-based approach to understanding children's early phonologies (e.g. Ingram 1974; Macken 1987; and see Macken 1995 for a review). Although most of these examined segmental phenomena, Judith Hochberg (1988a, b) examined the acquisition of stress by Spanish-speaking children.

One of the drawbacks of some of these early studies, however, was that many of the phonological processes proposed seemed unnatural, not being attested in the phonology of any language. The tools that generative phonology had to offer at the time were inadequate for dealing with many aspects of synchronic and diachronic phonological systems, as well as for processes of acquisition. Somewhat ahead of his time, David Stampe (1969) had proposed a perspective on developing phonologies and phonological process more generally involving the notion of constraints. Also emergent at the time were proposals for autosegmental representations, initially used to help handle tonal systems (Goldsmith 1976), but eventually applied to vowel harmony and other phonological systems as well. Andrew Spencer (1986) subsequently reanalyzed Neil Smith's (1973) acquisition data, showing how a more autosegmental approach could handle certain phonological processes that were unnatural from a linear approach to segmental phonology.

Although much of the early research on the acquisition of phonology focused on segments, some European researchers began to focus on the word as an important unit in children's early phonological organization. Drawing on insights from J. R. Firth (1948), Natalie Waterson (1971, 1987) proposed that children's early phonologies could best be characterized by a holistic, non-segmental-prosodic approach. These findings were followed by proposals by George Allen and Sarah Hawkins (1978, 1980) that English-speaking children's early words tended to take the form of disyllabic trochaic feet. They observed that children's early words are often augmented (CVC > CVCV) or truncated (e.g. *banana* > ['nænə]), both processes resulting in a disyllabic trochaic foot. They further proposed that such early

word shapes might be universal, representing the default, or unmarked form of early words.

Subsequent theoretical developments at the prosody-syntax interface (Selkirk 1984) coincided with further interest in examining higher-level structures in children's early speech as well. For example, Matthei (1989) investigated across-word process in children's early speech. Consistent with Allen and Hawkins (1978, 1980), he found that some lexical items were augmented to a disyllabic trochaic foot when produced in isolation (1a–b). However, when the two are combined into a larger phonological phrase, both were phonologically reduced (1c), again yielding a binary foot.

(1)		*Child*	*Adult target*		
	(a)	[ˈbebi]	/ˈbebi/	'baby'	(1;5)
	(b)	[ˈbʊkɔ]	/ˈbʊk/	'book'	
	(c)	[ˈbebʊ]	/ˈbebiz ˈbʌk/	'baby's book'	

These findings are extremely relevant for understanding the nature of phonological constraints on the shape of children's early words, especially given more recent developments in phonological theory (see Sections 5.1 and 5.3).

Around the same time, Marlys Macken (1978, 1979) began to observe templatic types of patterns in children's early words. Although somewhat in advance of John McCarthy's (1981, 1989) work on templatic phonology, the observations were similar. In particular, Macken noted that some children went through a period of development where their early words exhibited certain distributions of consonants, such as permitting CVCV sequences to contain consonants with only [+ant, −cor] features in onset position, and only [+ant, +cor] in medial position. Thus, words such as Spanish *Fernando* were realized as [mano], and *libro* 'book' as [pito].

Researchers had long known that the phonetic transcription of speech renders only a rough approximation of what was actually said. Although it is possible to conduct broad phonemic transcription with a relatively high level of consistency, the transcription of narrow phonetic detail (e.g. stress, voicing, vowel length, intonation contours) is not as reliable. This is even more problematic in the case of child speech. This presents a problem for understanding the nature of children's early phonological representations, and how they develop. For example, English-speaking children often appear to devoice final consonants (as in German). However, Stoel-Gammon and Buder (1999) note that the reliability of transcribing voicing in child speech is notoriously poor. Macken and Barton (1980a, b) found that, despite the apparent lack of a voicing distinction on stop onsets, both English- and Spanish-speaking children make an acoustic distinction in VOT between voiced and voiceless stops. Thus, a voicing contrast appears to be distinctive in English-speaking children, despite the fact that it may be unreliably perceived as such by adults. The finding of such *covert contrasts* has important implications for our understanding of the nature of feature representations in children's early phonologies, and points to the importance of conducting acoustic analysis for determining which feature contrasts children actually make. The

extent of covert contrasts in children's early speech is not entirely known. As noted in Section 5.2, this type of research is gaining ground again, due in part to a renewed interest in gestural approaches to phonology (e.g. Browman and Goldstein 1990) and acoustic approaches to feature cues (e.g. Stevens and Keyser 2010).

Drawing on Paul Kiparsky's (1973) developments in lexical phonology (as a way to constrain the ordering of rules on phonological representations), Kiparsky and Menn (1977) proposed a two-lexicon model for phonological acquisition: a perception representation and a production representation, and a process that mapped between them (see also Menn and Matthei 1992). More recently, Paul Smolensky (1996) has argued that positing two lexicons is not necessary, showing that the same ranking of constraints can be used in both perception and production, leading to well-formed lexical representations, but constrained output forms. Interestingly, Menn (1983), drawing on earlier discussions by Charles Kisseberth (1970) involving what he called *conspiracies*, also proposed that children's early phonologies were subject to output constraints. Thus, the seeds for a constraint-based approach to the acquisition of phonology, as well as phonology more generally, had been laid.

In sum, the early formulation of generative phonology, using segments and rules, was inadequate for handling many phonological processes, both in language, and in acquisition. This led to new developments in phonological theory, exploring non-linear, prosodic approaches to certain phonological processes. As shown below, it also led to a more empiricist cognitive/biological approach to understanding phonological development.

3.2 *Cognitive and Biological Approaches to Individual Differences*

One of the problems in the field of phonological acquisition has always been the lack of data. Many studies, like Neil Smith's (1973) study of Amahl, were individual case studies of only one child, limiting the ability to generalize about phonological development both within and across languages. One of the goals of the Stanford Child Phonology Project (1968–1988) was therefore to collect longitudinal audio recordings from several (e.g. five) children from different languages.

Ferguson and colleagues (Ferguson and Farwell 1975) conducted a large, cross-linguistic study (English, French, Japanese, Swedish, Ukrainian) examining children's transition from babbling to first words (birth to 18 months). The results showed that there was no silent period between the two (in contrast with what was proposed by Jakobson 1941). A number of researchers found that there was a close relationship between the sounds and sound sequences of babbling (at 6–8 months) and those used in first words (Cruttenden 1970; Menyuk 1968; Oller *et al.* 1976). It was also found that many of children's very first words appeared to be well formed, and then went through a period of reanalysis/generalization before once again becoming more adult-like (not unlike the U-shaped curve reported for morphological over-regularization) (e.g. Moskowitz 1973). However, although the first segments produced exhibited certain language-specific tendencies,

individual differences were also found for individuals learning the same language. Some of this took the form of phonological selectivity, where some children showed unexpected phonological processes (e.g. *snow* > [nos]). Ferguson and Farwell (1975) suggested that these individual differences were evidence of a cognitive strategy to language learning, where young children were "hypothesis testers," examining the articulatory space in an effort to produce certain segmental contrasts. This led to further investigation of phonological development in terms of a cognitive processes (e.g. Macken 1978, 1979; Macken and Ferguson 1981; Macken and Ferguson 1983; Menn 1971, 1976; Vihman, Ferguson, and Elbert 1986).

Although language-specific differences were found in children's early productions, Ferguson and Farwell (1975) also suggested that there were universal phonetic tendencies that were constrained by the biology of the human vocal tract and central nervous system. This raised questions about children's perceptual abilities and the relationship this had to their early language production. Locke (1983) proposed that biological models of language development might be more explanatory, more effectively handling the issue of continuity in language development. Kent (1984) was one of the first to think of early phonologies as self-organizing systems, where rhythmic structures were seen as a natural biological phenomenon that could account for some of Allen and Hawkins's (1979, 1980) observations about the prominence of trochaic feet in early grammars. MacNeilage and Davis (1990) further proposed that the observed prominence of CV syllable structures in children's early productions, as well as C-V interactions (e.g. high vowels more likely to follow coronal consonants, etc.), could best be understood in terms of constraints on mandibular oscillation. Others have incorporated ideas from dynamical systems theory, proposing that phonological development can best be explained with reference to increasing complexity of organization, or pattern formation adaptation, in terms of exploration of the sound space (Thelen 1981, 1989, 1991; Vihman 1993).

Björn Lindblom's (1983, 1992) concern with the learner's perceptual capacities, and the connection between perception and production, was consistent with his view that there was a tension in speech production between *information transfer* and *gestural economy*. That is, speakers try not to use more articulatory effort than needed to convey the intended information. The most stable representation for maximal clarity of acoustic cues is a series of CV sequences, which is also the most frequent syllabic sequence in children's canonical babbling and early speech (MacNeilage 1980; MacNeilage and Davis 1990). This may be a consequence of early lack of control of coarticulation, leading to the pervasive production of consonant harmony (Vihman 1978) and reduplicative forms in young children's early speech (e.g. Waterson 1971 *biscuit* [be:be:]). Recent developments in articulatory phonology (Browman and Goldstein 1990) have renewed interest in the role of articulatory gestures and their role in understanding phonological systems. As discussed in Section 5, these theoretical developments, as well as new (ultrasound) technology, now make it possible to explore the nature of articulatory constraints on children's early productions, and their relationship to phonological representations, in a non-invasive fashion.

In sum, the Stanford Child Phonology Project served as a major stimulus for the study of phonological acquisition beginning in the 1970s. Two conference proceedings volumes (Ferguson, Menn, and Stoel-Gammon 1992; Yeni-Komshian, Kavanagh, and Ferguson 1980) presented the state of the art at the time, and a reference point for researchers in a field that began to rapidly change in the 1990s. The Stanford studies were largely descriptive, providing for the first time developmental data from several children from five different languages from babbling until they had a vocabulary of approximately 50 words (around 1;6). However, it is around this point in development that evidence for children's phonological representations begins to be more robustly evidenced. Thus, although these studies provided ample evidence of early segmental development, and some word development, few testable predictions can be made about children's developing phonological representations using this data. Rather, what is needed is data spanning the time period from the onset of first words (around 1 year) until 2 or 2;6 years – data that, until recently, has been scarce. Also missing during this period was a theoretical framework from which to understand the nature of developing phonological representations. As discussed in the next section, both theory and data began to arrive in the 1990s, stimulating the field of phonology and phonological acquisition in new ways.

4 From the 1990s to 2007: Phonological Constraints

The 1990s had a significant impact on both the field of theoretical phonology and phonological acquisition. This was largely due to the stimulus provided by Optimality Theory (Prince and Smolensky 2004), which energized the entire field of phonological inquiry. However, the 1990s also saw the beginnings of more interdisciplinary research in phonological acquisition. One of the venues that provided a more theoretical stimulus to the field was the University of British Columbia's International Conference on Phonological Acquisition (Bernhardt, Gilbert, and Ingram 1996). This event brought together a wide range of European and North American speech-language and phonological researchers working on a wide variety of languages, stimulating new and continuing collaborations in the field. During this time researchers of language acquisition and infant-speech perception also began to interact, stimulated by an interdisciplinary conference at Brown University (Morgan and Demuth 1996). This was followed by a similarly inspired conference in Berlin, resulting in the volume by Jürgen Weissenborn and Barbara Höhle (2001). Both brought together researchers working on the perception and production of different languages, as well as more computationally oriented researchers, laying the groundwork for future collaborative endeavors. These interactions were reinforced by the broadening of the scope of papers accepted at the yearly Boston University Conference on Language Development to include aspects of infant-speech perception and computational approaches to learnability in addition to the more traditional theoretical linguistic treatments of language acquisition.

The 1990s therefore had a significant impact on the field of phonological acquisition, and this influence is ongoing today, with a growing international community of researchers working on the acquisition of phonology. This has come about, in part, through the development of a clearer set of predictions about the course of acquisition and how this might be realized in languages with different phonological structures. But it has only been possible to address these issues with the increased availability of longitudinal computerized corpora from more children between the ages of 1 and 3. These developments have led to a number of special conference sessions and special thematic volumes devoted to the acquisition of phonology, again providing an important reference for the training of future researchers. For example, Goad and Rose (2003) focused on segmental-prosodic interactions, with several papers discussing constraints on cluster reduction. René Kager, Joe Pater, and Wim Zonneveld (2004) included papers examining a variety of issues in phonological acquisition from an optimality-theoretic perspective. The papers in Demuth (2006) explore the impact of frequency effects on understanding cross-linguistic differences in prosodic word shape. All provide insight into the acquisition process as one of constraint satisfaction.

4.1 The Emergence of the Unmarked in Early Phonologies

As mentioned in Section 3, acquisition researchers had already experienced the limitations of rule-based, segmental accounts of children's early productions (Smith 1973), and had begun to explore other approaches to understanding the nature of early phonological systems. This continued in the 1990s. For example, Demuth (1993) used an autosegmental approach to the acquisition of Bantu tonal systems, showing that 2-year-old Sesotho speakers had no problem with lexical tone, but only acquired grammatical melodies that interacted with tone sandhi and OCP effects by the age of 3 (see Yavas 1994 and Archibald 1995 for other non-linear treatments of both first- and second-language acquisition). The development of a constraint-based approach to phonology (Prince and Smolensky 2004) provided further tools for exploring the nature of phonological acquisition.

As previously mentioned, one of the challenges for the field of phonological acquisition has always been the lack of data. It is relatively easy to conduct phonological and syntactic experiments with older children, and corpora available on the CHILDES database (MacWhinney 2000) include a rich source of information about children's syntactic structures. However, in 1990 there was little in the way of phonetically transcribed acquisition data available in the public domain. The publication of Paula Fikkert's (1994) thesis on the phonological development of 12 Dutch-speaking children, which took a parameter-setting approach to the acquisition of stress (Dresher and Kaye 1990), was therefore a significant event given the wealth of data it contained.

At the same time, Jane Fee (1995), and Katherine Demuth and Jane Fee were examining English-speaking children's early word shapes from a more prosodic perspective, trying to provide a unified explanation for both weak initial-syllable truncation and reduplication/vowel epenthesis. Drawing on developments in

prosodic phonology and morphology (Nespor and Vogel 1986; Selkirk 1984, 1996), they proposed that children's early productions are governed by highly-ranked No-Coda constraints, as well as constraints against initial unstressed syllables. They also suggested that children's early words were actually "minimal words," and that children's grammars provided support for the *emergence of the unmarked* (McCarthy and Prince 1994b), providing the first optimality-theoretic (OT) analysis of children's prosodic words; Demuth 1995). Demuth (1996) then reanalyzed Fikkert's Dutch data, showing that, like the English data, it could be handled from a constraint-based perspective, and that perhaps children learning all languages would exhibit a similar minimal word stage of early development. Gnanadesikan (2004) similarly proposed that the notion of the emergence of the unmarked could help account for children's early onset reductions at the level of the syllable. Pater (1997) then integrated these two proposals, showing that the children's early word truncations could be understood in terms of both higher-level prosodic constraints that also obeyed markedness constraints at the level of the syllable. Thus, the truncation of *banana* to ['bænə] preserved an obstruent at the beginning of the syllable, the unmarked option for onsets.

Since that time, there have been numerous OT analyses of children's developing phonologies, and even those that do not provide a formal OT analysis are often inspired by the notion of constraints. Some of this was anticipated by phonologists such as Stampe (1969), who, like Kisseberth (1970) and Paradis (1988), saw phonological systems as a constraint-satisfaction problem rather than as a set of phonological rules. This perspective has provided the field of phonological acquisition with an extremely useful framework from which to explore the nature of children's developing prosodic phonologies. We provide the highlights of some of these studies below.

4.2 The Acquisition of Syllable Structure

Although Erik Fudge (1969) had pointed to the importance of the syllable as a unit of phonological analysis, this was a relatively neglected area of research in the generative program until the work of G. N. Clements and Samuel Keyser (1983) (see Chapter 5, this volume). The development of notions relating to the sonority hierarchy and the sonority sequencing principle (e.g. Clements 1990; Ladefoged 1993) set the stage for examining further phonotactic restrictions on syllable structure acquisition. For example, Kehoe and Stoel-Gammon (2001) found that alveolar stops were the first coda consonants to be acquired in English. This was of significant interest since sonorants are typically assumed to be the unmarked form for coda consonants crosslinguistically, exhibiting less of a sonority rise between nucleus and coda than from onset to nucleus. However, Stites *et al.* (2004) showed that alveolar stops are the most frequent coda consonants in English, and confirmed that children tend to acquire these first, rather than the less frequent, phonologically unmarked sonorant codas. Thus, although frequency and markedness typically pattern together, most children may show a preference for frequency over markedness effects in their early productions, all else being equal. This raises

questions about the notion of markedness as a whole, and its relationship to frequency for learners of a particular language. It also raises the question of which linguistic units learners are using for calculating frequency. For example, Zamuner *et al.* (2004) show that coda consonant production is a function of neighborhood density. That is, it is the frequency of the rhyme+coda, rather than simply the coda consonant itself, that is the best predictor of accuracy in coda consonant production, at least English. On the other hand, /ʁ/ is one of the most frequent consonants in French, yet several studies have found that at least some French-speaking children have persistent problems with the production of /ʁ/ (e.g. Demuth and McCullough 2009a; dos Santos 2007; Rose 2000). This may be due to articulatory problems with this uvular fricative, or due to its variable realization in the input children hear.

Developments in feature geometry (e.g. Clements 1985) and gestural approaches to phonology (Browman and Goldstein 1990) also provided the theoretical background for examining consonant-vowel interactions within the syllable. For example, Claartje Levelt (1995) found that Dutch-learning children were much more likely to produce onset consonants that shared the features of the following vowel. Thus, the consonants preceding back rounded vowels tend to be labials, and those preceding high vowels tend to be coronals. This raises the question of when and how learners begin to be able to represent cues to feature contrasts in an adult-like fashion. This will be explored further in Section 5.

Research on the structure of the syllable has also given rise to many studies on the acquisition of more complex syllable structures. Much of the early work on consonant clusters was carried out with children who exhibited phonological delay, raising questions about the representation of clusters in children's early phonologies more generally (e.g. Chin and Dinnsen 1992; Gierut 1999; see Bernhardt and Stemberger 1998 for a review). Some of this research focused on the factors governing cluster reduction, with different proposals as to the constraints involved. Following Pater (1997), many researchers have proposed that children will typically preserve the least marked onset, i.e. the least sonorant segment of the cluster (e.g. Barlow 1997; Diane Ohala 1996, 1999). Thus, in the case of the word *stop*, the obstruent /t/ would be preserved, but in the case of the word *sleep*, the /s/ would be preserved.

However, others have noted the limitations of the sonority account (e.g. Barlow 1997, 2001; Goad and Rose 2004; Pater and Barlow 2003). Goad and Rose (2004) proposed that children preserve the consonant that is the head of the syllable. Thus, in an obstruent + liquid cluster such as *plate*, the obstruent /p/ is the head of the syllable and will be preserved. In contrast, the /s/ in s-clusters is syllabified as an adjunct, leaving the /l/ as the preserved head of the syllable in a word like *slate*. However, Pater and Barlow (2003) show that one child simplified the onset of *sneeze* to /n/, whereas the onset to *sleep* was simplified to /s/. Jongstra (2003) therefore proposed that when the sonority distance is close, the segment contiguous with the nucleus will be preserved, whereas when the sonority distance is sufficiently far, the least sonorous segment will be preserved. A recent study of cluster simplification calls all the above into question, noting that features from

both consonants often remain in cluster reduction (Kirk 2008). Interestingly, the majority of studies of consonant cluster reduction have examined the acquisition of Germanic languages. Little is known about consonant cluster reduction processes in other languages.

All of the above mentioned studies examined onset clusters. There has been little research on the acquisition of word-final clusters. One might expect these to be acquired later due to the fact that codas are typically considered more marked than onsets. However, Lleó and Prinz (1996), in a longitudinal study of five German-speaking 1–2 year-olds, found that word-final clusters were acquired several months earlier than word-initial clusters. Furthermore, Levelt *et al.* (2000) found that nine of the children in the Dutch CLPF corpus first acquired clusters word-finally, whereas only three of the children first acquired clusters word-initially. They suggest that this variability in Dutch patterns of syllable structure development is due to the fact that the frequency of consonant clusters in both positions is approximately the same. Controlling for segments, Kirk and Demuth (2005) found that English-speaking 2-year-olds were also more accurate at producing word-final as opposed to word-initial consonant clusters. Overall, coda clusters are more frequent than onset clusters in English. However, the children also exhibited better production of final nasal+s and stop+s clusters than final nasal+stop and s+stop clusters. Furthermore, children often metathesize the s+stop clusters (*wasp* > *waps*), suggesting that frequency or articulatory factors may be involved. Note also that the most accurately produced clusters are those that typically occur with morphologically complex forms, suggesting that morphology may provide a perceptual or production advantage for some consonant clusters.

To control for the possible effects of frequency and morphology, Demuth and Kehoe (2006) conducted an elicited production study to examine the acquisition of consonant clusters in French. The found that 2 year-olds were more accurate at producing onset rather than word-final clusters, a finding confirmed in a subsequent longitudinal study (Demuth and McCullough 2009a). Some researchers have proposed that some word-final consonants in French (and other languages) prosodify as onsets to empty-headed syllables (e.g. *partir* 'to leave' /paʁ.ti.ʁØ/) (Charette 1991). It is possible that this structure is more marked, and therefore acquired later (though Goad and Brannen 2003 claim that such structures are universal at early stages of acquisition). Rose (2000) noted, however, that one child from his longitudinal study of two children learning Canadian French had acquired all but /ʁ/ in word-final position, but had /ʁ/ as a coda word-internally. He therefore proposed that this child had a coda representation for /ʁ/ in all positions. However, researchers have also reported that the acoustic and articulatory characteristics of French /ʁ/ are extremely variable, both between speakers and within the speech of individual speakers (see Demuth and McCullough 2009a for review). Little is known about the acquisition of segments that are variably realized in the input, or where the syllabic representation is potentially ambiguous (though see discussion of the acquisition of branching onsets and rising diphthongs in Rose (2000) and Kehoe *et al.* 2008).

4.3 The Acquisition of Prosodic Word Structure

The early work on the acquisition of prosodic word structure (Demuth 1995; Pater 1997) suggested that children had an early awareness of word-minimality effects, and that this could be captured in terms of constraint interactions. Further support for this proposal came from Mitsuhiko Ota (1999). Using acoustic analysis, he showed that young Japanese-speakers who cannot yet produce coda consonants exhibit compensatory lengthening of the vowel, thereby preserving bimoraic (and minimal word) structure. In many respects, this study was ahead of its time, providing acoustic/phonetic evidence to support the theoretical claims made.

One of the issues raised by the Word Minimality Hypothesis was what predictions it would make for the acquisition of a language such as French, where a large portion of the lexicon violates word-minimality. Demuth and Johnson (2003) examined this issue in longitudinal data from one French-speaking child. They found that the child's earliest words (1;3–1;5) were all target or reduplicated CVCV forms. As in other languages, her early grammar showed a highly-ranked constraint against word-final (coda) consonants, resulting in either reduplicated CVCV repairs, or truncated CV outputs. Thus, for certain CVC target words she produced subminimal, monomoraic words. More striking, however, was the reduction of disyllabic CVCV words to monosyllabic CV form. Further analysis showed that segmental constraints against fricatives, velar stops, and clusters were more highly ranked than faithfulness to syllable preservation and/or word minimality (see dos Santos 2007 for similar observations from another child who does have velar consonants). French is also a language that permits subminimal, monomoraic CV lexical items, and these constitute approximately 20% of all words French-speaking children hear (Demuth and Johnson 2003). In keeping with similar proposals by Levelt *et al.* (2000), the authors suggest that learners are sensitive to the high-frequency phonological structures of the ambient language, and adjust their grammars (constraint ranking) accordingly. This can be understood in terms of trying to be as faithful as possible to the input forms, thereby minimizing constraint violations. Note that such a perspective on the development of early grammars minimizes the role of universal markedness. Rather, higher-frequency phonological forms become the "unmarked" structures on a language-specific basis.

On the other hand, Goad and Buckley (2006) proposed that French learners do show early word-minimality effects. They report that a Canadian child showed compensatory vowel lengthening when the word-final consonant was missing, though no supporting acoustic analysis was provided. However, further analysis of spontaneous productions from two children from France showed no systematic lengthening of the vowel when the word-final consonant was missing (Demuth and Tremblay 2008). The number of subjects examined in all these studies was small, suggesting that further study with more children at the early stages of acquisition (1–2 years) is needed to resolve this issue.

Demuth *et al.* (2006) returned to the issue of word-minimality in English, drawing on new data from four children between the ages of 1–3. Although some

children showed apparent compensatory vowel lengthening (and one child showed an early period of epenthesis, where CVC words surfaced as CVCV), this occurred on both monosyllabic and disyllabic words, and on both long/tense as well as short/lax vowels. If learners were using compensatory lengthening to preserve word-minimality, one would expect this to be restricted to monosyllabic targets with short/lax vowels, the context where a second mora of structure is required to preserve a bimoraic foot, or minimal word. Further acoustic analysis of three children's compensatory processes found that two of the children exhibited across-the-board compensatory lengthening for missing codas, whereas only one (older) child showed compensatory lengthening only for target words with a short/lax vowel (Song and Demuth 2008). These findings suggest that children may initially exhibit compensatory lengthening for omitted coda segments in English, and only later (around the age of two) come to realize that languages like English observe word-minimality constraints.

Note that the implications of the English findings on compensatory vowel lengthening contrast with those of Mitsuhiko Ota (1999), where he proposed that early vowel lengthening to compensate for missing codas in Japanese provided support for an early awareness of moraic structure. There are three possible explanations for this. First, Japanese is a mora-timed language, and children may become more aware of moraic structure and its consequences for prosodic word structure earlier in such a language. Second, although Japanese, like French, permits subminimal words, Ota reports that Japanese parents, in speaking to their young children, generally augment subminimal forms. This raises the issue of the nature of the input children hear (child-directed speech), and the effects this may have on children's developing phonologies. Finally, it is possible that both Japanese- and English-speakers compensate for missing segments at the ends of words. However, since coda consonants are always moraic in Japanese, it is difficult to determine if compensatory lengthening is due to segmental versus moraic factors. This is obviously an area for further cross-linguistic research.

Roark and Demuth (2000) proposed that the frequency distribution of syllable and prosodic word structures in the input children hear could influence the types of prosodic word structures children used in their early utterances. In a corpus analysis of child-directed speech they showed that the token frequencies of word shapes in English and Spanish are significantly different. In particular, the majority of words in English are monosyllabic, and there are very few trisyllabic words. This contrasts with Spanish, which has many more trisyllabic and tetrasyllabic words. They suggested that this different distribution of word shapes could account for English-speaking children's tendency to truncate words like *banana* until around the age of 2;6 (Pater 1997). In contrast, Spanish-speaking children appear to permit prosodic words of larger structure earlier than English-speaking children. Several of the papers in Demuth (2006) pursue this issue further, bringing a much-needed cross-linguistic perspective on these issues. Vigário, Freitas, and Frota (2006) and Lleó (2006), analyzing acquisition data from European Portuguese and Spanish respectively, found that the relative frequency of word shapes in the input helps explain the truncation patterns. This is consistent with the findings from English,

Dutch, and French mentioned above. However, Prieto (2006), in a comparative analysis of the acquisition of Spanish and Catalan, suggested that the relative frequency of foot shape, rather than prosodic word shape, helps explain why Catalan learners (but not Spanish learners) pass through a stage of development where they truncate disyllabic S(w) prosodic words. Finally, Ota (2006) suggested that lexical frequency, rather than prosodic word shape, best accounted for the few cases of truncation found in children's Japanese. Thus, frequency effects at various levels of structure may help determine constraint rankings in the grammars of children learning different languages, resulting in different truncation patterns in early prosodic word development.

Importantly, these patterns of truncation appear to be due to phonological, not perceptual or articulatory constraints. Interestingly, Carter and Gerken (2004) found that, when children omitted the initial unstressed syllable of a three-syllable word following a verb, they left a prosodic trace of the missing syllable, which is realized as a silent interval. This suggests that, at least in some cases of syllable omission, children have in some sense planned for the syllable, even though no segmental content is realized. Such findings provide further evidence of covert contrasts in children's early speech that are often missed in traditional phonetic transcription. These findings raise questions about the extent to which other apparent omissions in early child speech may in fact be prosodically, gesturally, or acoustically realized at some level of analysis. They also point to the need for a developmental model of speech planning and production. Both would provide a better understanding of the phonological representations and phonological constraints that govern children's early phonological grammars.

4.4 *Acquisition at the Phonology/Morphology Interface*

Drawing on insights from the Prosodic Hierarchy, and proposals for the prosodification of grammatical morphemes, researchers in the 1990s also began to examine the nature of prosodic constraints on children's realization of grammatical morphemes. Early research on southern Bantu languages (which exhibit penultimate lengthening at the end of a phonological phrase) had noted that children tended to produce noun class prefixes with monosyllabic stems, but omit them with disyllabic stems (e.g. Connelly 1984). Demuth (1994) proposed a phonological explanation of this issue, observing that children tended to produce noun class prefixes when these constituted part of a disyllabic foot (*mo-tho* 'person'), but were more likely to omit them when they were unfooted (*mo-sadi* > [sadi] 'woman'), much like the truncation of prosodically similar English monomorphemic words (e.g. *banana* > ['nænə]). This phenomenon begins to disappear around the age of 2;3, once children's prosodic phonologies develop further (Demuth and Ellis 2009). LouAnn Gerken and colleagues (Gerken 1994; Gerken and McIntosh 1993) also noted that English learners were more likely to produce grammatical morphemes such as pronouns and determiners when these could be prosodified as part of a foot. In the meantime, Selkirk (1996) proposed a set of constraints on prosodic well-formedness, where prosodic structures immediately dominated by the next

higher level of structure within the Prosodic Hierarchy conformed to the constraint on exhaustive prosodic parsing, or Exhaustivity. Thus, prosodic clitics that could not be prosodified as part of a foot would violate Exhaustivity. Gerken (1996) capitalized on this development, proposing that children would be more likely to include grammatical morphemes in early speech if these could be prosodified as part of a foot, thereby conforming with the constraint on Exhaustivity. However, unfooted grammatical morphemes would need to be prosodified at a higher level of structure (e.g. the phonological phrase), and would therefore violate the Exhaustivity constraint. Thus, children's variable omission of grammatical function items could be understood in terms of markedness, where those determiners that incurred the fewest constraint violations would appear earlier in children's productions.

Lleó (1996) had long noted that Spanish-speaking children (unlike German-speaking children) exhibit the use of proto-determiners from the beginning of their speech. This was explained in terms of the high frequency of Spanish three-syllable words, which provide children below the age of 2 with a prosodic window large enough to permit the prosodic licensing of determiners (Demuth 2001; Lleó 2001; Lleó and Demuth 1999). Support for this hypothesis came from the fact that three-syllable words that were truncated to two syllables are nonetheless accompanied by a (proto)determiner (e.g. *la muñeca* 'the doll' > [a'meka] (Demuth 2001, Demuth *et al.* in press)). This suggests that Spanish-speaking children have an early three-syllable prosodic word window that can be adapted to permit inclusion of a determiner plus a following foot.

Research on other languages shows a similar tendency for children to produce prosodically licensed grammatical morphemes earlier than those that are not, thereby accounting for some of the variable production of grammatical function items in children's early speech. For example, Demuth and Tremblay (2008), in a study of two French-speaking children, showed that determiners begin to robustly appear with monosyllabic words around 1;10 years, but only begin to be consistently used with disyllabic words a few months later. The authors suggest that the early determiners are prosodified as part of an iambic foot. Only once children begin to represent higher levels of prosodic structure do they begin to include determiners with longer words, showing evidence of having acquired adult-like prosodic structure, where determiners are prosodified at the level of the phonological phrase.

Demuth and McCullough (2009b) found a comparable pattern of longitudinal development for the acquisition of articles by five English-speaking children. Similar to the French findings, and consistent with the experimental findings from Gerken (1996), they show that four of the children had significantly higher use of articles when these could be prosodified as part of a foot with the preceding word. In contrast, the children tended to omit the articles that remained unfooted (i.e. those prosodified at the level of the phonological phrase) (e.g. *Tom [hit the]*$_{FT}$ *ball* vs. *Tom [patted]*$_{FT}$ *(the) ball*). Furthermore, this pattern persisted for 4–5 months, disappearing as the children reached the age or 2–2;6 years. This is about the same time that children begin to more reliably produce the initial unstressed syllables of lexical items like *banana* (cf. Pater 1997).

These findings suggest that children's acquisition of grammatical morphemes is closely tied to the development of prosodic representations. Given that many grammatical morphemes are unstressed prosodic clitics, their acquisition is dependent on the development of higher-level prosodic structures. The Prosodic Licensing Hypothesis therefore provides a framework for exploring the development of higher-level prosodic representations, and how this changes over time. It also provides a principled means for making predictions about the course of grammatical morpheme development within and across languages. As shown in the case of Spanish determiner acquisition, the development of these constraints on prosodic structure is also closely tied to the prosodic properties of the lexicon, though the exact relationship between the two is not yet clear.

The prosodic licensing of grammatical morphemes appears to occur at the level of the syllable as well (e.g. Stemberger and Bernhardt 1997). Recent research has found that there are phonotactic and positional effects on the acquisition of English third-person *-s*. That is, children are much more likely to produce this grammatical morpheme when it forms a simple coda than when it forms part of a consonant cluster (e.g. *sees* vs. *hits*), and when it appears utterance finally compared to utterance medially (Song *et al.* 2009). These findings suggest that there is still much to be discovered about the phonology-syntax interface in children's developing grammars, where constraints on prosodic representations may account for much of the variable production of grammatical morphemes.

4.5 *Featural Underspecification and Phonological Processes*

With the development of approaches to feature underspecification (Archangeli 1988; see Steriade 1995 for a review) came a renewed interest in trying to understand segmental phenomena in children's early grammars. For example, Stemberger and Stoel-Gammon (1991) and Stoel-Gammon and Stemberger (1994) proposed that consonant harmony processes, which are common in early English, could be understood in terms of underspecification theory. Goad (1997) took this further, proposing that consonant harmony in children's early grammars could be characterized in terms of the relative ranking of parse/link place features and those that align place features at the left edge of the harmonic domain (see dos Santos 2007 for discussion of similar phenomena in French). Others explored which theory of underspecification (radical, contrastive) could best account for children's phonemic inventory, and how this changed over time (e.g. Gierut 1996). Morrisette *et al.* (2003) explored issues of markedness and the representation of place features, making the prediction that if the child's system represented dorsal place distinctions, it would also represent coronal place. Furthermore, they proposed that coronals would be expected to replace dorsals, but not vice versa. Interestingly, this prediction is not upheld in the cases of consonant harmony in one French-speaking child, where dorsals regularly replace coronal consonants that are not in the head of the foot (dos Santos 2007).

4.6 Frequency Effects and Phonological Acquisition

Every since Noam Chomsky (1965) defined the field of formal linguistics as constituting the knowledge of grammar (linguistic competence, or I-language), issues of language performance (language use, E-language) have been largely ignored, or relegated to the field of sociolinguistics. However, researchers have long been aware that issues of lexical frequency, for example, play a large role in understanding aspects of grammaticalization and historical change, many of which involve phonological and morphophonological processes (e.g. Bybee 2001, 2007). Psycholinguists have also long known that adults appear to have encoded not only information about the phonology, morphology, syntax, and semantics of lexical items, but also the likelihood that a given lexical item will appear in a given grammatical context (e.g. MacDonald *et al.* 1994). Furthermore, this notion of *predictability* is encoded in how we speak, with high-frequency and/or predictable information typically being phonologically (or at least phonetically), reduced (e.g. Aylett and Turk 2004; Lindblom 1983). That is, the speaker appears to have a model of the listener in mind, and will phonetically reduce redundant or less important information when speaking.

Recent research on infant-speech perception has shown that infants are extremely sensitive to the frequency of segments and prosodic structures in the primary language to which they are exposed (e.g. Anderson *et al.* 2003). Thus, despite the fact that there is some noise in the signal, language learners appear to be capable of extracting the information needed to create both a phonology and a lexicon (see Saffran *et al.* 1996) for discussion of whether these abilities are domain specific, or more general). It has also long been known that 3–5-year-old children's morphophonological representation of familiar, high-frequency words is more robust in both perception and production than that of novel and low-frequency words (Edwards *et al.* 2004). More recently, scholars have begun to examine more systematically the effects of frequency on the acquisition of syllable and prosodic word structures (e.g. Zamuner *et al.* 2004).

Recall that Roark and Demuth (2000), in a corpus analysis of child-directed speech, found that English had a much higher frequency of syllable final (coda) consonants than did Spanish (60% vs. 25%). These frequencies closely matched those suggested by Pierre Delattre (1965) for adult speech. They used this to help explain the earlier acquisition of coda consonants in English relative to Spanish. In a similar vein, Levelt *et al.* (2000) found that the frequency of different syllable structure shapes (e.g. CVC, etc.) in Dutch corresponded closely to the order in which these were acquired. Importantly, they also showed that the fit was much better using a corpus of child-directed speech rather than adult-directed speech, indicating that the course of acquisition adheres closely to the statistics of the lexicon children typically hear (cf. Ota 2006). Furthermore, they found that, when the frequency of syllable structures was the same (CCVC, CVCC) nine of the children acquired the less marked complexity in the onset first, and only three acquired the more marked complexity in the coda first.

These results, showing that language learners have an early sensitivity to the frequency of different phonological and prosodic structures, is interesting in light of proposals that children would first acquire crosslinguistically unmarked structures (e.g. Demuth 1995; Gnanadesikan 2004; Jakobson 1948; Pater 1997). Although unmarked structures are often extremely frequent, this is not always the case. For example, given the restriction found in many languages where only sonorant consonants are permitted in the coda (e.g. Japanese), it is often assumed that sonorants are the unmarked form for coda consonants (cf. Clements and Keyser 1983). However, in languages like English, the most frequent coda consonants are stops, the most frequent of these being /t/. That is, the highest frequency coda consonants in English are the most marked crosslinguistically, raising questions about the order of acquisition of coda consonants in this language. To investigate this issue, Stites *et al*. (2004) conducted a longitudinal study of three children, investigating which coda consonants were the first to appear. They found that two of the children showed the frequency pattern, whereas only one showed the markedness pattern. Cross-sectional findings by Kehoe and Stoel-Gammon (2001) confirm that /t/ is the first coda consonant to appear. Thus, it appears that, for most children, robust frequency effects will typically override more cross-linguistic (i.e. markedness) tendencies, at least in some prosodic domains.

One of the challenges to any study of frequency effects is to determine what to count. The results mentioned above indicate that language learners may be keeping statistics over any number of different types of linguistic units simultaneously. In the case of phonology, this might include every level of the Prosodic Hierarchy, and segmental interactions therein. For example, much of the research on lexical acquisition finds that children's accuracy in the production of lexical items is closely related to neighborhood density (Edwards, Beckman, and Munson 2004; Storkel 2004). Furthermore, the acquisition of coda consonants appears to be closely linked not only to coda consonant frequency, but also to neighborhood density within the entire rhyme (Zamuner *et al.* 2004). Thus, some of the variability found in the acquisition of syllable structures, as well as words and morphemes, may be explained by the frequency with which these are segmentally and phonotactically represented in the lexicon.

Issues of lexical frequency, as well as acoustic and articulatory factors, may account for some of the variable production of coda consonants in different prosodic contexts. For example, Demuth *et al.* (2006) showed that English-speaking children acquire word-final coda consonants earlier in monosyllabic as compared with disyllabic words. In an experimental study with novel disyllabic words, Kirk and Demuth (2006) showed that children were more likely to produce the same coda consonant when it occurred either in a stressed syllable or at the end of the word. They suggest that this is due to the longer duration found in both positions, providing increased time for the inclusion of another articulatory gesture. Thus, frequency effects appear to help explain some of the variance in phonological acquisition within a certain class of prosodic structures. But across prosodic contexts, either within or across languages, other contextual and/or gestural planning phenomena may better account for some of the variable production found. Thus,

although the frequency of coda consonants is much higher in English compared to Spanish, the earlier acquisition of coda consonants in English may also be due to the fact that most English lexical items used in everyday speech are stressed monosyllables. This contrasts with Spanish, where coda consonants often fall on unstressed syllables and/or at the ends of polysyllabic words (e.g. *escaleras* /es̲.ka.ˈle.ras̲/ 'stairs'). Thus, in addition to frequency effects, prosodic factors such as position within the word and phrase (Hsieh, Leonard, and Swanson 1999), as well as stress, may also play an important role in determining the nature of children's early syllable, word and morpheme productions. Such issues are not currently incorporated into models of early acquisition, in part due to the focus on linguistic competence rather than on performance factors. However, most of the child language data come from performance of some kind, be it perception, comprehension, or production (though often not recognized, this is true of adult linguistic data as well). By controlling for the factors outlined above, it may be possible to better understand the nature of children's knowledge of phonology (linguistic competence), and when various phonological structures have been acquired.

5 The Future of Phonological Acquisition

The field of phonological acquisition has grown significantly since the early 1990s. This has been largely due to theoretical developments in phonology, combined with the new availability of phonologically transcribed longitudinal data from several children in different languages. New developments are also now taking place in experimental methods, acoustic analysis, and computational modeling, pointing to a vibrant future for the field. This section highlights some of these developments and their theoretical import.

5.1 Constraint-based Approaches

The field of phonological acquisition has been significantly influenced by the development of constraint-based approaches to the study of phonological systems (e.g. Prince and Smolensky 2004). This has provided a framework for investigating interactions between different types of constraints in the developing system, and for viewing phonological acquisition as a constraint-satisfaction problem. That is, given the limited processing/production capacities of a 1–2-year-old, how does the child make himself or herself understood? Viewed as a constraint-satisfaction issue, it is possible to see that there are several different solutions to a given phonological problem, giving rise to certain types of well-attested individual variation. This approach therefore provides a framework for understanding individual variation in terms of differently ranked (or weighted) constraints, much in the way that dialectal variation occurs in closely related languages. Thus, although Optimality Theory and related approaches provide little in the way of predictions about what is possible or impossible in early acquisition (this

is presumably governed by constraints on universal grammar as well as the developing physiology of the child), it does provide a framework for exploring the nature of constraint interactions for a given child learning a given language. What determines the particular ranking of constraints, and how this changes over time, will presumably be determined, in part, by the frequency of different constraint violations in conjunction with general markedness factors. Thus, given an initially highly-ranked constraint against coda consonants (No-Coda), the learner of a language with many codas (e.g. German or English), will typically demote this constraint faster (to avoid massive constraint violations) than the learner of a language with fewer codas (e.g. Spanish) – all else being equal (e.g. Boersma and Levelt 2000). Furthermore, some of the variability found within a given child, for the same target word in the same sentential context, can be handled in much the same way that variability has been handled in adult phonological systems (e.g. Anttila 2002; Nagy and Reynolds 1997), that is, in terms of floating or overlapping constraints (Boersma and Levelt 2000; Demuth 1997).

As in the field of phonology itself, each of these issues was intractable within the framework of a rule-based system. A constraint-based perspective therefore provides a means for better describing what is happening during the acquisition process. This in turn lays the groundwork for developing an explanation for the process, and making predictions about how a given phonological phenomenon will be acquired in other domains or other languages. To address these issues we will need access to new data sources, something that has already begun to develop (see Section 5.3).

5.2 *Articulatory/Acoustic Approaches*

At the same time that constraint-based perspectives on phonology began to grow, articulatory approaches to phonology also began to develop (Browman and Goldstein 1990). If feature contrasts are the outcome of articulatory gestures, then surely this has implications for understanding the nature of early phonological acquisition as well. Although this perspective on language acquisition has been slower to develop, it is beginning to be investigated more seriously in cases of disordered speech. For example, using ultrasound technology, Bernhardt *et al.* (2005) have found that English-speakers who exhibited persistent problems with the production of /ɹ/ were using only one of the articulatory gestures needed to produce this segment. Once they were given intervention on the appropriate gesture to use, they quickly began to pronounce /ɹ/ as appropriate to a North American west coast dialect. Thus, some of the constraints on children's acquisition of phonology may be articulatory, especially in cases where multiple articulatory gestures are required (e.g. liquids, clusters, affricates). This is obviously an area for future research.

Given the complexities of articulation, there may also be acoustic and/or articulatory evidence that children are actually approximating certain feature contrasts, despite the fact that these are often not perceived as such by the listener/transcriber. In the past few years there has been renewed interest in investigating

such covert contrasts (e.g. Scobbie *et al.* 2000), and we expect such lines of inquiry to continue, providing a richer set of acoustic evidence for children's developing phonological representations. Critical to such an endeavor would be the investigation of how children represent certain acoustic landmarks in their early speech productions, and how they enhance certain gestures or acoustic features to ensure that certain feature contrasts are clear (e.g. Keyser and Stevens 2006; Stevens and Keyser 2010). Stoel-Gammon and Buder (2002) showed that most English-speaking children control extrinsic vowel lengthening before voiced/voiceless consonants by the age of 2. However, Demuth *et al.* (2006) also suggest that those children who exhibit a period of word-final vowel epenthesis on CVC targets, may do so to ensure that the voicing cues to the final consonants are clear (see Section 4.3). Further research has shown that many of the acoustic cues to early voicing contrasts (voice bar, coda release) are already adult-like by the age of 1;6 (Shattuck-Hufnagel *et al.*, forthcoming).

The acoustic analysis of children's early speech may provide evidence for their developing phonological representations in other domains as well. Recall that Song and Demuth (2008) found compensatory lengthening on CVC targets when the word-final consonant was omitted (CVC > CV:). In a very different domain, acoustic analysis of English showed that the one child stressed her articles at early stages of acquisition, with a decrease in interval length between the article and the preceding word over time (Demuth and McCullough 2009b). Thus, by the age of 2, this child was finally beginning to treat articles as prosodic clitics rather than as independent prosodic words. Little is known about the prosodic organization of children's early productions, and how this interacts with both prosodic constraints and planning/production issues. Using acoustic and articulatory information (including the use of ultrasound measures) may help to address these issues.

5.3 New Longitudinal Databases and Data Analysis Tools

One of the challenges for the field of phonological acquisition has been the lack of publicly available longitudinal phonetically transcribed data from multiple children between the ages of 1 and 2 years, even for well-studied languages like English. The availability of this type of data is particularly important due to the fact that children are actively acquiring segmental, syllabic, and prosodic word structures during this time, as well as early morphology. This lack of early spontaneous production data has been especially problematic since it is often difficult to conduct elicited production experiments with children below the age of 2, and it is extremely time consuming to collect and transcribe large samples of spontaneous speech.

Beginning with the collection of semi-longitudinal semi-cross-sectional data from 12 Dutch-speaking children (the CLPF database), there is now an increasing number of phonetically transcribed longitudinal corpora from different languages. Many contain child speech from the onset of first words until around the age of 3, as well as interactions with parents, providing some indication of the nature of the input these children receive. Some of these databases also contain

linked acoustic files, facilitating verification of the phonemic transcription, and allowing for the acoustic analysis of both child and adult speech. Fortunately, many of these databases are now publicly available on CHILDES (MacWhinney 2000). A summary of some of these resources is presented in (2) (see http://childes.psy.cmu.edu/ for the latest updates and documentation/manuals).

(2) Early longitudinal data containing IPA transcription

Language	*Children*	*Ages*	*Hours*	*Database*
Catalan	4	1–4	112	Serra-Solé Catalan Corpus
Dutch	12	1–3	132	CLPF Corpus
English (American)	6	1–3	365	Providence Corpus
English (British)	2	1–3	30	Cruttenden Corpus
English (British)	1	2–3		Smith 1973 Corpus
English (Canadian)	2	2–4	30	Montréal English Corpus
French (France)	4	1–3	185	Lyon Corpus
French (Canadian)	2	2–4	30	Québec French Corpus
Japanese	3	1–2	75	Ota Corpus
Portuguese (European)	8	1–4	140	Freitas Corpus
Spanish (Spain)	1	1–3	40	Llinàs-Ojea Corpus
Spanish (Spain)	1	1–3	50	López Ornat Corpus

The collection and transcription of child speech corpora is an extremely time-consuming and labor-intensive task. The increasing availability of this type of data in the public domain will facilitate the investigation of phonological development for years to come. However, the number of children included in these studies is still very limited. Thus, there will be an ongoing need for other data sets in the future, designed to address specific phonological research questions. Although data collection and archiving, as well as the availability of Unicode IPA fonts, make this process easier than ever before, automatic transcription of the acoustic signal (forced alignment) has yet to be perfected, even for adult speech. However, in addition to scripted programs, the tool PHON is now available for conducting phonological analysis (Rose *et al.* 2006), and is freely downloadable from the CHILDES website. Acoustic analysis can also be carried out using downloadable Praat tools (Boersma and Weenink 2005). All of these developments should facilitate the collection and analysis of additional phonological acquisition corpora in the years to come, providing critical information about the initial stages of phonological acquisition, and the implications for the emergence of grammatical morphemes.

5.4 New Experimental Methods

Spontaneous production corpora can provide a wealth of information about children's acquisition of phonological units that occur frequently in spontaneous

speech (e.g. segments, syllable structures, and word truncation). However, it is less informative regarding the acquisition of lower frequency phonological phenomena (e.g. certain grammatical morphemes, cluster types, or word shapes). Elicited production experiments are therefore helpful in providing sufficient tokens of the phonological issue being investigated while also controlling for issues of segmental and prosodic context (though they typically cannot be used below the age of 1;6). Many of the studies discussed above have employed various elicited production techniques with both familiar and novel words, examining the constraints on children's early language productions using perceptual and acoustic measures. New methods using ultrasound technology are also beginning to examine the nature of children's articulatory gestures between the ages of 1–4 (Gick 2007; Ménard, Loevenbruck, and Savariaux 2006). These acoustic and articulatory studies hold much promise for better understanding the nature of children's early phonological representations, and how these develop over time.

Infant-speech perception studies can also be used to investigate learners' early sensitivities to different types of phonological structure (see Morgan and Demuth 1996 and Jusczyk 1997 for reviews). It has long been known that infants can discriminate between native and non-native segmental contrasts by 11 months (e.g. Best, McRoberts, and Sithole 1988). Recent studies show that, by 19 months, infants have highly sophisticated featural representations that are sensitive to changes in voicing, manner, and place, at least for consonants at the beginnings of familiar words (White and Morgan 2008). Furthermore, a series of studies has recently shown that, for fourteen-month-olds, recognizing new words presents more of a challenge than recognizing known words (cf. Swingley 2007 for a review). Infants also show a very early preference for listening to the prosodic structure of their native language (Nazzi and Ramus 2003), and can pick out high-frequency familiar words (such as their name) from the speech signal by 6 months (Bortfeld *et al.* 2005). However, identifying disyllabic words with iambic structure in the speech stream presents a challenge until 16 months for infants learning both English and French (Nazzi *et al.* 2006). Taken together, these findings suggest that the phonological representation of familiar words is more robust than that of words that have only recently been encountered.

As the infant studies have become more sophisticated, moving from segments to words and morphemes, the fields of production and perception are beginning to overlap (Gerken and McIntosh 1993). Some of the first studies that test both perception and production in the same children in a referential task indicate that the connection between the two might be tighter than often assumed (Sundara *et al.* 2011). This is obviously an area for further research.

5.5 Modeling Phonological Learnability

Along with developments in OT came a renewed interest in addressing phonological learnability issues. Some of the first research to examine this issue took the perspective of constraint reranking as a function of accumulating constraint violations (Prince and Tesar 2004; Tesar and Smolensky 2000). Other research has

explored constraint reranking in terms of frequency-induced changes in constraint weights (Albright and Hayes 2006; Boersma and Hayes 2001; Boersma and Levelt 2000). Recent models have begun to explore more probabilistic approaches to learning constraint rankings (Goldwater and Johnson 2003), phonological categories (Goldsmith and Xanthos 2009), and syllable structures (Goldwater and Johnson 2005; Hayes and Wilson 2008). As in any modeling enterprise, the question is always what type of input the machine learner receives, and how representative this is of what human learners have to work with. This issue is especially relevant for modeling the acquisition of phonology, where the machine learner is often trained on segmented words, whereas the human learner must identify words from the unsegmented acoustic representation of the speech stream. Thus, there is still much to be done in terms of developing more psycholinguistically plausible models of phonological learning.

6 Conclusion

The field of phonological acquisition has grown significantly since the development Optimality Theory. This has provided a much-needed framework for exploring the nature of the constraints on early phonological grammars, and how these change over time. It has also coincided with an increase in the number of phonetically transcribed corpora from different languages containing longitudinal spontaneous production data from children between the ages of 1–3. This has begun to allow the field to make and test predictions about the factors that influence the acquisition of phonology, and how this develops over time. In conjunction with new experimental methods investigating perceptual, acoustic, and gestural aspects of phonological development, and more sophisticated means of modeling the learning process, the future of the field is open-ended. New findings should help provide a clearer picture of how phonological systems are acquired, with implications for better understanding the nature of phonological disorders, the evolution of language, phonological change, and possibly phonological theory itself.

18 Phonology as Computation

JOHN COLEMAN

1 Transformational Generative Phonology and the Sequential Machine Architecture

In its earliest years, generative phonological theory was strongly influenced by work in computer science, theoretical and applied. In the Cold War climate of the 1950s, the US Army (Signal Corps), the US Air Force (Office of Scientific Research, Air Research and Development Command), and the US Navy (Office of Naval Research) supported research by Harvard University's George Miller and Roman Jakobson, Morris Halle, Kenneth Stevens, Noam Chomsky, Yehoshua Bar-Hillel and colleagues in the Acoustics Laboratory and the Research Laboratory of Electronics at MIT, on a variety of problems of speech processing and computational linguistics, with a particular focus on Russian and English. Their work was informed by and contributed to research in information theory, automata theory, and logic.

One influential line of inquiry at that time concerned the application of Shannon's mathematical theory of communication (Shannon 1948) to linguistics and psychology. Despite Chomsky's trenchant and influential arguments against the probabilistic approach, work such as Straus (1950), Cherry, Halle, and Jakobson (1953),[1] Mandelbrot (1961), and Miller and Chomsky (1963) illustrates the extent to which information theory was taken seriously, in the tradition of phonological statistical analysis exemplified by Trubetzkoy (1939: Part I, Chapter 7), Saporta (1955), Hockett (1955: 138–143, 215–218), and Harris (1955); the influence of Harris and Jakobson on the early work of Chomsky and Halle, respectively, is well documented elsewhere.

Another line of inquiry concerned the nature of the most appropriate data structures for encoding phonological information. Fully recognizing that, as an

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

acoustic phenomenon, speech is continuously varying, Halle (1959a: 504) defended a representation in terms of a string of discrete segments as follows:

> While a rigorous segmentation procedure which would show in all cases a one-to-one correspondence with the linguistic representation may not be possible, it is possible to construct devices which produce speech by utilizing a set of discrete instructions which coincide closely with the linguistic segmentation. The devices I have in mind are of the type of the Bell Telephone Laboratories' Voder or the Haskins Laboratories' Octopus [*two early speech synthesizers – JC*]. The signal emitted by these devices is continuous speech, yet the input instructions are discrete.

Each unit in the string of discrete segments was in turn broken down into a number of distinctive features, so that utterances as a whole were represented in terms of a rectangular matrix of segments → features, such as Table 18.1.

Not content with merely representing information, the early generative phonologists employed two related metaphors for how linguistic forms might be generated: (i) actual digital computers, then quite new, and (ii) the derivation of theorems from axioms in formal logic. Halle (1959b: 12–13) puts it thus:

> an adequate description of a language can take the form of a set of rules – analogous perhaps to a program of an electronic computing machine – which when provided with further special instructions, could in principle produce all and only well-formed (grammatical) utterances in the language in question. . . .
>
> Like all other parts of the grammar a phonological description is formulated here as a set of deductive rules.

The same commitment to explicitness in specifying the "program" for a language's phonology is maintained in Chomsky and Halle (1968: 60):

> The rules of the grammar operate in a mechanical fashion; one may think of them as instructions that might be given to a mindless robot, incapable of exercising any judgment or imagination in their application. Any ambiguity or inexplicitness in the statement of rules must in principle be eliminated, since the receiver of the instructions is assumed to be incapable of using intelligence to fill in gaps or to correct errors. To the extent that the rules do not meet this standard of explicitness and precision, they fail to express the linguistic facts.

It is instructive to compare this to the following passage from Rosenbloom (1950: 160), an advanced logic textbook cited by Chomsky (1956):

> Another way of looking at these languages is to consider the productions as instructions to a moron, who can scan a string and recognise it as being of a certain form, for producing theorems starting from the axioms. The happy moron can, by merely following the instructions, generate as many theorems as he pleases, and never feels the need for any intelligence in the process. He might just as well be a robot or machine.

These lines immediately precede an exposition of Post's production systems, a powerful approach to the specification of the formal languages of logic and

Table 18.1 An early example of a feature matrix (Jakobson *et al*. 1952), an "analytical transcription" of the phonologically rich sentence "Joe took father's shoe bench out." Reprinted with permission from *Preliminaries to Speech Analysis: The Distinctive Features* and their Correlates, by Roman Jakobson, C. Gunar, M. Fant and Morris Halle, published by MIT Press.

	ʒ̂	ˈo	u	t	ˈu	k	f	ˈa	ə	ð	ə	z	ʃ	ˈu	u	b	ˈe	n	ʃ̂	#	ˈa	u	t
Vocalic	(–)	+	+	(–)	+	(–)	(–)	+	+	(–)	+	(–)	(–)	+	+	(–)	+	(–)	(–)		+	+	(–)
Compact	+	+	–	(–)	–	+	(–)	+	–	(–)	–	(–)	+	–	–	(–)	+	(–)	+		+	–	(–)
Grave		+	+	–	+		+	±	±	–	±	–		+	+	+	–	–			±	+	–
Nasal	(–)			(–)		(–)	(–)			(–)		(–)	(–)			(–)		+	(–)				(–)
Tense	–			+		+	+			–		–	+			–			+	(–)			+
Optimal Constrictive	±			–		–	+			±		+	+			–			±				–
Stressed		+	–		+			+	–		–			+	–		+				+	–	

mathematics with important similarities to transformational grammars (as noted by Scholtz and Pullum 2007: 718). There is then the following prophetic note:

> A *canonical* language is a language *L* with a finite alphabet, a finite number of productions, and a finite number of axioms. . . . One might also expect that many concepts in linguistics which have resisted all attempts up to now at a clear and general formulation may now be treated with the same lucidity and rigor which has made mathematics a model for the other sciences.

While Halle (1959b) expressed the rules of Russian phonology in ordinary English, the algebraic form of the proposed rules was soon settled upon (see Halle 1962): they were to be transformational rules that is, rewriting rules of a relatively unrestricted type that transform one sequence of segments (each a vector of distinctive features) into some other. All rules were of the type "A → B in the environment C — D," in which A is a single symbol. Although the changes from A to B were in practice minimal and phonetically reasonable, in principle the formalism permits arbitrary transformations of a string into any other.

Such grammars were most obviously interpreted as devices for *generating* an observable phonetic output from a hypothesized stored lexical form. Nevertheless, Halle and Stevens (1959) had also conceived of generative phonology being employed in an *analysis-by-synthesis* scheme to derive linguistic events from the acoustic level. They observe:

> The analysis procedure that has enjoyed the widest acceptance postulates that the listener first segments the utterance and then identifies the individual segments with particular phonemes. *No analysis scheme based on this principle has ever been successfully implemented.* (Emphasis added)

Their generative alternative (see also Bell *et al*. 1961) is:

> a recognition model in which mapping from signal to message space is accomplished largely through an active or feedback process. Patterns are generated internally in the analyzer according to a flexible or adaptable sequence of instructions until a best match with the input signal is obtained. Since the analysis is achieved through active internal synthesis of comparison signals, the procedure has been called *analysis by synthesis*.

Thus generative phonology, instantiated using a transformational grammar in conjunction with phonetic-to-acoustic mapping rules and an analysis-by-synthesis scheme, could be seen as a computational model of the processes of speech production and recognition.

By the early 1970s, the application of generative phonology to computer programs for speech synthesis and recognition was being explored. Carlson and Granström (1974) presented "a phonetically oriented programming language" for synthesis of speech by rule, and Cohen and Mercer (1974) presented a similar rule system for use in an automatic speech-recognition system. For unavoidable

engineering reasons (i.e. because they did not work very well), such rule systems were not successful for automatic speech recognition. Rule systems flourished for a while in text-to-speech systems, most notably MITalk (Allen, Hunnicutt, and Klatt 1987), though not in quite the form proposed by phonological theorists; and because of the laboriousness of discovering and writing the phonological rules for a language, the "hand-crafted" approach to synthesis-by-rule has been largely abandoned in favor of automated, data-oriented approaches.

To illustrate the current variety of approaches to phonological computation, we shall consider some specific cases from Spanish that exemplify and test each framework's account of syllabification and phonotactics. In particular:

> *Syllabification:* For input strings such as /tiempo/ and /buei/, how many syllables are there? Which segments are nucleus peaks? Where do the vowels /i/ and /u/ surface as a syllable nucleus and where as a glide?
>
> *Phonotactics:* Why are initial */ml/ and */sl/ ill-formed in Spanish, whereas initial /pl/ and /fl/ are well-formed? How is the language-specific nature of these facts accounted for?

Early transformational generative phonology could address these phonotactic questions in two ways. First, a morpheme structure rule (e.g. Halle 1959b: 56–62) might specifically prohibit certain combinations of segments. Alternatively, combinations of segments may be systematically absent at the output of the grammar because of the operation of some rule; as an (implausible but illustrative) example, a grammar including the rule "s → f / — l," systematically turns /sl/ info /fl/, accounting for the absence of /sl/ and occurrence of /fl/. By either means, however, the "explanation" of which combinations do or do not occur is language specific because the rules of early transformational generative phonology are language specific.

Chomsky and Halle (1968) did not address the sorts of syllabification problem set out above,[2] but soon phonologists began proposing rules to insert syllable boundaries (e.g. Vennemann 1971; Hoard 1971). Hooper (1972) examines some Spanish data concerning nasal+obstruent clusters for which it proves advantageous for rules to refer to syllable boundaries. Rule (1) puts a syllable boundary between two syllabic segments in such Spanish forms as *se*$*a* "let it be" and *di*$*a* "day":

(1) Ø → $ / [+syllabic] — [+syllabic]

Rule (2) states that if there is only one non-syllabic segment between two syllabic segments, the $-boundary occurs before the non-syllabic segment. This rule ensures that the non-syllabic segment is a syllable onset, not a coda (though no such constituent categories are overtly employed in this analysis): *u*$*na* 'one', *o*$*so* 'bear', *ha*$*ya* 'let there be', *pe*$*ro* 'but'.

(2) Ø → $ / [+syllabic] — [–syllabic] [+syllabic]

When there are two or more non-syllabic segments, if one is an obstruent other than /s/, and the following segment is a liquid or glide, the syllable boundary occurs before the obstruent: *pa$dre* 'father', *con$tra* 'against', *si$glo* 'century', *res$plandor* 'splendor', *a$byerto* 'open', *a$bwelo* 'grandfather'. If there are two non-sonorant segments, the $-boundary is inserted between them: *ap$to* 'apt'. Except for the restriction concerning /s/, this is (part of) a fairly general, cross-linguistic rule of rising sonority in onset consonant clusters:

(3) $\emptyset \rightarrow \$ \;/\; [+\text{syllabic}]\ [-\text{syllabic}]_0 \;—\; [-\text{sonorant}]\ [+\text{sonorant}, -\text{nasal}]_0\ [+\text{syllabic}]$

There are several exceptions to rule (3). First, coronal obstruents do not occur syllable initially before /l/, that is, /tl/ and /dl/ are divided between two syllables:

(4) $\emptyset \rightarrow \$ \;/\; [-\text{sonorant}, +\text{coronal}] \;—\; [+\text{lateral}]$

Rule (3) will place a $-boundary between an /s/ and a following obstruent if the cluster is flanked by vowels: /Vs$tV/, as in *es$tar*, *es$pecial*, but it incorrectly places a $-boundary *before* the cluster /sr/ or /sl/; in Spanish the syllabification should be /s$r/ or /s$l/, as in *Is$rael*, *is$la* "island." Thus Hooper also proposes the following:

(5) $\emptyset \rightarrow \$ \;/\; [-\text{sonorant}, +\text{coronal}, +\text{continuant}, +\text{strident}] \;—\; [+\text{consonantal}, +\text{vocalic}]$

The language-specific nature of rules (4) and (5) vs. the claimed universality of the others is not a difference that is explicitly captured.

Finally, if there are two nasals, two liquids, liquid plus nasal or nasal plus liquid, a $-boundary separates them: *hon$ra* 'honor', *per$la*, 'pearl', *al$ma* 'soul'. If there is one consonantal segment, it begins the syllable: *a$lyento* 'breath'. For these, Hooper gives this rule, which is rather similar to (3):

(6) $\emptyset \rightarrow \$ \;/\; [+\text{syllabic}]\ [-\text{syllabic}]_0 \;—\; [+\text{consonantal}]\ [-\text{consonantal}]_0\ [+\text{syllabic}]$

When these rules are applied in the order given (though this is not critical as they are disjunctively ordered) to the five test items, the derivations are as follows:

Rule no.	*buei*	*tiempo*	* *mleta*	* *sleta*	*plata*	
(1)	*bu$e$i*	*ti$empo*	–	–	–	
(2)	–	–	*mle$ta*	*sle$ta*	*pla$ta*	
(3)	–	*tiempo*	–	–	–	cf. *u$na $plata*
(4)	–	–	–	–	–	cf. ? *d$leta*
(5)	–	–	–	*sleta*	–	cf. *esleta*
(6)	–	–	–	–	–	
Outputs:	*bu$e$i*	*tiempo*	? *mle$ta*	? *sleta*	*pla$ta*	cf. ? *u$na m$le$ta*

There are two inadequacies with these outputs. First, *buey* /buei/ and *tiempo* come out wrongly, because of hiatus rule (1). To prevent this, /w/ and /j/ could be encoded as glides underlyingly. But then, rather than *predicting* whether vocoids are syllabic or non-syllabic, the input would state whether they are [±syllabic]; in some other approaches this is not necessary. Second, *mleta*, *sleta*, *dleta*, which are supposed to be ill-formed, undergo syllabification without any difficulties; these rules do not tell us why the outputs *s*$*le*$*ta* and *d*$*le*$*ta* are ill-formed. Presumably, some prohibition against, for example, /s/ or /d/ occurring as an independent syllable is needed, such as *$[–syllabic]$, but transformational generative phonology does not provide for such constraints on outputs.

Transformational generative phonology's combination of language-specific rules, rule ordering, derivations, and abstract underlying representations drew (and continues to draw) criticism on a variety of fronts. During the 1970s, generative phonologists attempted to address some of these concerns. Principles of rule ordering were sought (see Anderson 1974) and non-derivational statements regarding the phonological relatedness of words were introduced ('via rules': Hooper 1976). A number of proposals regarding phonological representations (Anderson and Jones 1974; Goldsmith 1976; Liberman and Prince 1977) led to some simplification and systematization of phonological rule types, and the introduction of other formal devices for defining grammatical representations, such as well-formedness constraints.

At the same time, concerns about the excessive computational power of the transformational grammar formalism mounted, especially in syntax and mathematical linguistics. Chomsky (1965: 208, Note 37) had claimed that languages generated by transformational grammars formed a proper subset of the recursive sets, and that they can thus be recognized effectively, but Peters and Ritchie (1973) put paid to this belief by showing that the transformational grammars of Chomsky (1965) generate all recursively enumerable languages, including many undecidable languages that could not be possible human languages (see Levelt 1976 and Bach and Marsh 1978 for introductory explanations of the problem; Wasow 1978 discusses its consequences). One response to these technical difficulties (e.g. Chomsky 1981) was to retain movement transformations, devoid of language-specific qualifications, and to replace rules by an interacting system of cross-linguistic principles and parameters. Other syntacticians, notably Kaplan and Bresnan (1982) and Gazdar (1982), proposed the complete elimination of transformations, necessitating a variety of less powerful alternative mechanisms instead. The warm reception that such ideas met was bolstered by their more-or-less successful adoption in large-scale computational systems for parsing, machine translation and so on (e.g. Alshawi 1992; Grover, Carroll, and Briscoe 1993).

Naturally, some phonologists also began to seek alternatives to the transformational toolkit. There was at the same time a growing recognition of the importance of "output-oriented," apparently teleological constraints on phonological patterns, an oft-cited example of which was identified by Kisseberth (1970): in a number of American languages, various rules of vowel deletion and insertion conspire to ensure that clusters of three consonants are avoided. The only phonological patterns

or templates in transformational rules, however, are the structural descriptions or "inputs" to rules, and all transformations expressed derivational *changes*. What seemed to be required as well, or instead, was direct description of the well-formed and prohibited surface structures, via constraints such as "avoid X," "X is ill-formed," "a well-formed X consists of a Y and a Z," and so on. Such non-derivational constraints were already common currency in syntactic theory. For example, Chomsky and Lasnik (1977) employed surface filters to solve certain problems of overgeneration. To prevent sentences with two complementizers, such as *I saw the man [*who that*] you talked to, they proposed the "Doubly Filled Comp Filter," expressed:

(7) $*[_{\text{COMP}}\ X^{\text{max}}\ \text{complementizer}]_{\text{COMP}}$.

The same sort of device could easily be used to prohibit unattested phonotactic combinations, for example:

(8) $*[_{\text{ONSET}}\ C[+\text{labial}]\ w]_{\text{ONSET}}$
or, for Kisseberth's examples, perhaps: *CCC.

Post-transformational syntactic theories made extensive use of combinatorial constraints on syntactic features, prompting a detailed examination of the formal basis of feature theories by Gazdar *et al.* (1985: Chapter 2), who modeled syntactic statements such as [+INV] ⊃ [+AUX, FIN] upon phonological marking conventions such as [+nasal] → [+sonorant].

Some phonologists adopted Gazdar *et al.*'s strategies for eliminating transformations and the distinction between underlying and surface representations from phonological theory (see Bird and Ellison 1994; Bird and Klein 1994; Scobbie, Coleman, and Bird 1996; Coleman 1998: Chapter 5). Others retained a distinction between at least two levels of phonological representation (i.e. lexical vs. surface): *n*-level phonology (instantiated as "Cognitive Phonology" by Lakoff 1993, "Harmonic Phonology" by Goldsmith 1993b and "Two-Level Phonology" by Karttunen 1993, Kaplan and Kay 1981 [1994]) and Optimality Theory (Prince and Smolensky 2004) follow this research path. In the following sections, we shall unpack the computational bases of these approaches.

2 Finite State Approaches

When 1960s transformational grammar was found to be excessively powerful, syntacticians and phonologists investigated what restrictions on the formalism are necessary and sufficient, seeking a formalism that was restricted as possible, but not over-much. Non-transformational syntacticians came to a consensus that some mildly context-sensitive formalism is needed, but as phonology lacks the recursive hierarchical structures and long-distance dependencies encountered in syntax, it appeared likely that a strictly finite-state formalism would be adequate.

Computational linguists like finite-state formalisms because they are well-understood, efficiently computable, and correspond with the reality of all actually-existing computers, both electronic and neurophysiological (i.e. human brains), as opposed to abstract computing devices such as Turing machines, which hypothetically may have infinite memory or an unlimited number of processors. Non-finite-state formalisms, such as context-free grammars, context-sensitive grammars, or transformational grammars, cannot be faithfully implemented in any existing or conceivable real-world computer: in order to use them for real-world computations, it is necessary to impose resource constraints such as finite memory or processor time. Such formalisms, when thus trimmed down, are then no more powerful than finite-state.

Finite-state methods had been used in studies of language prior to generative grammar by, for example, Markov (1913), Shannon (1948), and programmatically by Hockett (1955, 1958: 291), but soon came to be strongly criticized as inadequate for natural language – including phonology – by Chomsky (1956) and Halle (1962), criticisms which were widely accepted and became enshrined as linguistics doctrine until the recent renaissance of interest in such methods (see e.g. Pereira 2000). Unfortunately, Chomsky's transformational grammars proved to be untenable, as Peters and Ritchie showed. In phonology, things are just as bad: even though phonological rules are context-sensitive in form, such grammars are more powerful than context-sensitive if they include deletion rules (Coleman 1998: 80–83), as they often do. This means that given a surface string, there is no guarantee that the underlying lexical items can be determined, whether by inversion of the rules or analysis by synthesis (see also Bear 1990). Such unrestricted transformational grammars are formally equivalent to Turing machines, an abstract, theoretical device capable of computing *any* computable function. Although it is nice sometimes to have powerful computing resources for one's work, it is of little help to either the working linguist or to a child trying to figure out the rules of its language to use grammars that are so excessively unconstrained. We hope to find constraints on the rules of grammar in order to explain how the infant learner discovers a good grammar of any language relatively rapidly. For this reason, we would like the *most* constrained, *least* powerful phonological formalism that is adequate to describe any language.

Fortunately, Johnson (1972) discovered that the way in which phonologists actually use rewriting rules rarely if ever exploits the full potential power of the transformational formalism.[3] He showed that, provided two conditions are observed, a set of rules of the form "A → B in the environment C — D" has no greater power than that of a regular grammar or finite-state machine, the least powerful of the Chomsky hierarchy of formal languages and their associated abstract automata. The two conditions are (i) that features must have a fixed range of values (e.g. + or –, or 1/2/3, but not unbounded integers), and (ii) rules must apply in a fixed order and may not reapply to their own output. These conditions are slightly problematic for phonological theory, but not unattainable. In practice, only the [stress] feature had unbounded integer values in Chomsky and Halle (1968): this representation of stress was replaced by metrical trees or grids, following

Liberman and Prince (1977), thereby satisfying Johnson's first requirement. The second condition is more difficult to satisfy, as it prohibits cyclic rule application. However, the range of phenomena attributed to cyclic rules has steadily reduced, calling their need into question (Cole and Coleman 1993). Inspired by Johnson's result, Kaplan and Kay (1981 [1994]) showed how phonological rules could be encoded as finite-state transducers; Koskenniemi (1983) developed a similar two-level treatment of morphophonological relations. Since a collection of finite-state transducers can be combined into one, a system of phonological rules can be compiled into a single (possibly quite large) finite-state transducer.

The words or expressions of any regular language can be generated or recognized by a finite-state automaton. These have a mathematical specification, described in most formal language theory textbooks, and can be implemented as working software. An automaton can be depicted as a network of nodes, each representing a *state* of the automaton, and arrows, representing changes from one state to another. The arrows are labeled with sets of symbols, one of which must be read or written by the machine in order to move from one state to the next.[4] Figure 18.1 shows an automaton which can be used to generate or recognize many one- and two-syllable words of Spanish.[5]

The automaton of Figure 18.1 can be used to generate a word as follows. Node 1, dashed, is the start state. From here we can move to state 2, 3, 4, or 5, by following an arrow. In doing so, one of the phoneme symbols on the arrow should be written down. For example, to move from state 1 to 4, the letter 'p' may be written. From state 4, a move to state 5 is permitted only if an 'l' is written. A path can be followed from state 1 to state 8, 16, or 17, writing out a sequence of symbols as we proceed, for example, 'p, l, a, n, t, a' or 'b, u, e, i'. States 8, 16, and 17 are end-states, shown by a bold circle.

A partial encoding of Figure 18.1 amenable to computational implementation is Table 18.2, a symbol–state table. Starting in state 1 (the second row), select any of the numbered cells in the row, write out the symbol at the top of the column, and move to one of the states (rows) given in the cell. Dashed cells show which symbols may not be written in that state. Cells containing more than one number have several possible next states, making this automaton *non-deterministic* when used for generation: at this point, some mechanism would have to guess or choose which state to go to next. Non-deterministic choices in an automaton lead to ambiguity, that is, to multiple possible analyses of a given input. For example, if a string beginning "piks . . ." is presented to the automaton in Figure 18.1, it may follow states 1, 5, 6, 7, 9 . . . or the different sequence 1, 5, 6, 9, 13. . . . There is a general method to make non-deterministic automata deterministic, if required (Hopcroft, Motwani, and Ullman 2001: 60–64), but we might prefer to keep the non-determinism. Although systems of phonological rules are usually written so that they can be followed deterministically, I know of no evidence from linguistics, psycholinguistics, laboratory phonology or wherever as to whether phonological processing by humans is deterministic or non-deterministic. The main motivations for using non-deterministic automata in computational linguistics are (a) compactness of the automata, and (b) modeling multiple possible "next moves" in processing

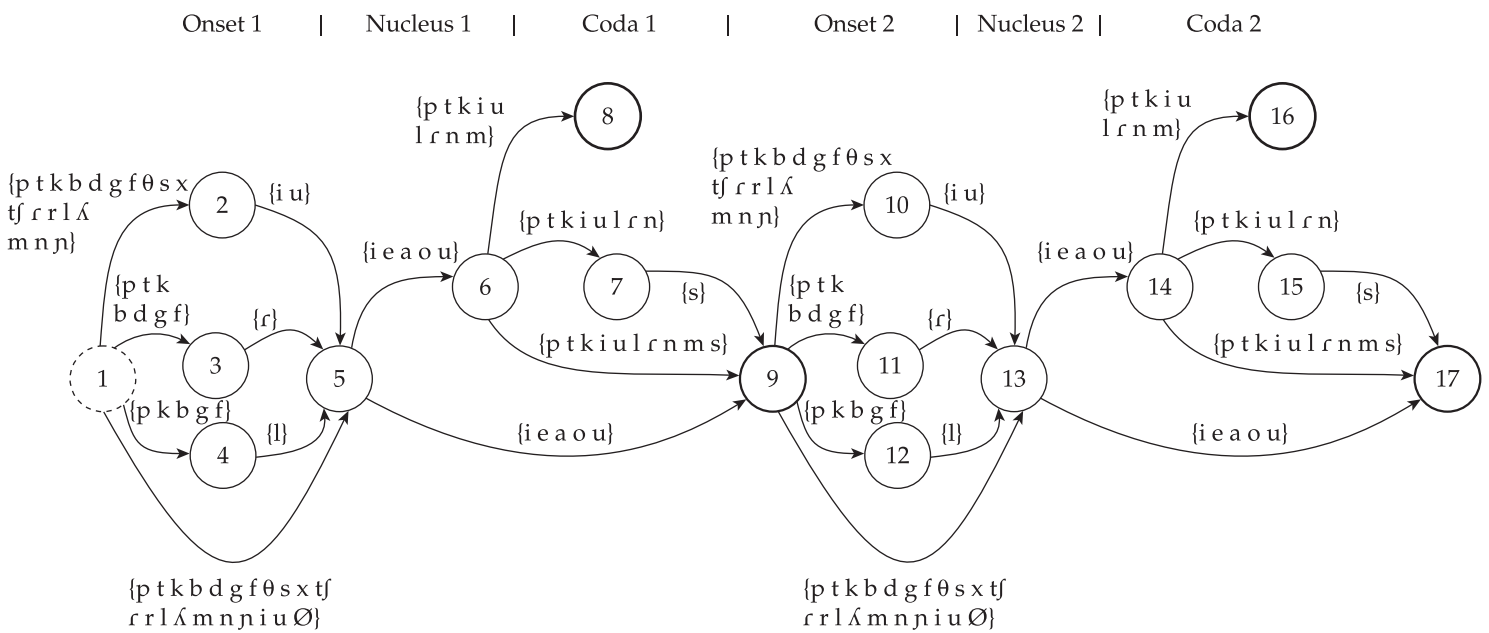


Figure 18.1 An automaton which defines many of the syllables and disyllables of Spanish.

Table 18.2 Symbol–state table encoding part of the automaton in Table 18.1.

Symbols:	b	d	f	g	i	k	l	ʎ	m
State									
1	2, 3, 4, 5	2, 3, 5	2, 3, 4, 5	2, 3, 4, 5	–	2, 3, 4, 5	2, 5	2, 5	2, 5
2	–	–	–	–	5	–	–	–	–
3	–	–	–	–	–	–	–	–	–
4	…	…							
⋮									

Symbols:	n	ɲ	p	ɾ	r	s	t	tʃ	θ	u	x
State											
1	2, 5	2, 5	2, 3, 4, 5	2, 5	2, 5	2, 5	2, 3, 5	2, 5	2, 5	–	2, 5
2	–	–	–	–	–	–	–	–	–	5	–
3	–	–	–	5	–	–	–	–	–	–	–
4											
⋮											

a string, as in the multiple choice networks used in Systemic Phonology (e.g. Halliday 1992). The ability to generate phonotactic ambiguity is desirable: for example, consider English *di.sco.ver* (with /sk/ the onset of the second syllable) vs. *dis.co.lor* (with /s/ and /k/ in separate syllables; or *Ritz* (final /ts/ a syllable coda) vs. *writs* (with, in some analyses, /s/ a syllable- or word-level appendix). As the generation of multiple possible analyses is an essential prerequisite of the "generate and test" method employed in constraint-based approaches (e.g. Optimality Theory's combination of Gen + Con), it appears that there has been a deeper change in phonological theory, abandoning determinism and embracing non-determinism.

Finite state automata can also *recognize* strings,[6] if the symbols on each arrow are read in and checked, rather than written out. For instance, the string "plata" is acceptable to the automaton of Figure 18.1: "p" takes the automaton from state 1 to either 2, 3, 4, or 5, "l" from 4 to 5 (the only possible way forwards), and so on via states 9, and 13 to 17. But 'mleta' (or any string beginning with *ml*) is unacceptable to Figure 18.1 (unrecognizable): 'm' takes the automaton from state 1 to state 2 or 5, but further moves are impossible, as none of the transitions out of states 2 or 5 are labeled with 'l'. Thus, *mleta* is not a well-formed Spanish word.

2.1 Finite State Transducers

Paired phonemic and phonetic symbols enable us to generate or recognize the specific details of pronunciation at different positions in a word. Thus, many aspects of pronunciation variation according to context may be modeled. Finite state

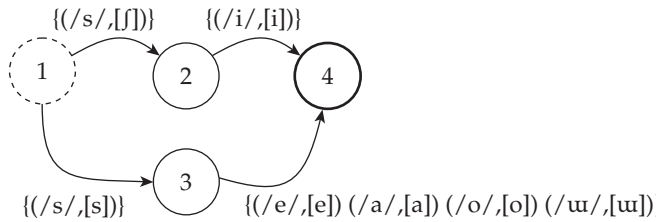


Figure 18.2 A finite-state transducer which encodes the phonological rule of Japanese that /s/ is pronounced as [ʃ] when it occurs before /i/, usually formalized as /s/ → [ʃ] — /i/.

transducers (i.e. automata with paired labels) formally implement two-level phonological rule systems (Kaplan and Kay 1981 [1994]), as most phonological rules may be expressed as finite-state transducers. For example, the Japanese rule that /s/ is pronounced as [ʃ] when it occurs before /i/ is encoded as a transducer in Figure 18.2. An entire rule system can be translated into finite-state transducers and then automatically combined into a single, large automaton, effectively eliminating rule ordering. Karttunen (1987) sketched the application of this method to an example of putative rule ordering in Klamath from Halle and Clements (1983); for details see Bear (1990), who also gives a two-level analysis of some phonological details of Tunica. Lakoff (1993) shows that a three-level version of this approach is sufficient to model other complex problem sets from Halle and Clements (1983), including some thorny problems from Mohawk and Yawelmani phonology.

The transducer in Figure 18.2 illustrates how *replacement* rules of the form "A → B in the environment C — D" are handled using finite-state transducers; we say that the *correspondence* A:B (in this case, s:ʃ) labels the transition from state *m* to state *n*, where C:something labels the transition(s) into state *m* and D:something labels the transition(s) out of state *n*. In environments other than C — D, a different state sequence is followed, in which (e.g. lexical) A corresponds to (e.g. observed) A, that is, A remains "unchanged."

Insertion rules of the form "Ø → X in the environment C — D" (epenthesis rules) are handled in the same way, except that instead of correspondence label A:B we must say that the empty symbol corresponds to X. Likewise, deletion of X is re-cast as a correspondence between X and the empty symbol, in some context.[7]

Extensions of these finite-state methods have been developed to deal with more complex phonological representations than strings of letters. Kay (1987) proposed a method for generation and/or recognition of the non-concatenative morphology of Arabic, in which vowel and consonant sequences form separate morphemes that must be intercalated according to particular patterns of C and V "slots." In Kay's finite-state treatment (following the autosegmental analysis of McCarthy 1979), the CV pattern, the sequence of vowels and sequence of consonants are written onto (or read from) three separate, parallel "tapes" under the control of a finite-state machine with a read/write head that looks at the three separate tapes simultaneously. A fourth tape is also employed: it stores the surface form

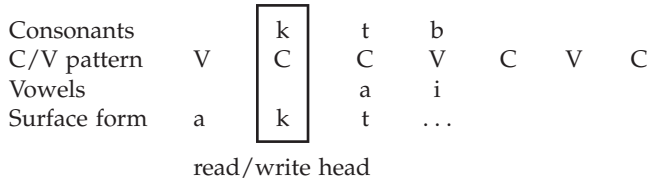


Figure 18.3

of the word being processed (Figure 18.3). Though Kay presents this as a way of handling nonconcatenative morphology, it is obvious that the content of the parallel tiers is computationally immaterial, and that the method could be used for processing any kind of autosegmental representation. For example, the three tiers could just as well be for segments, moras, and tones. Kay's proposal is informal; the nub of the idea was taken up and formalized in slightly different ways by Wiebe (1992), Pulman and Hepple (1993), Bird and Ellison (1994), and others.

There are many ways in which suprasegmental structure may be included in such devices. First, we might regard certain regions of a finite-state machine, that is, certain groups of states and transitions, as corresponding to suprasegmental constituents. For example, the portion of Figure 18.1 from state 1 to state 5 defines the onset of the first syllable, from state 5 to states 6 or 9 delineates the first nucleus, from 6 to 9 the first coda. If there is a second syllable, the portion from states 9 to 13 – essentially identical to the portion from states 1 to 5 – defines the second syllable onset, and so on. I have written the names of these constituents at the top of the figure, above the relevant regions; these constituent labels form no part of the formal specification of the automaton, but are added for expository purposes.[8] This is the approach to higher-level structure in finite-state automata taken by Carson-Berndsen (1998); it has the potential advantage of enabling us to be noncommittal about the suprasegmental status of units at the junction of two domains. In the English word *stupid*, /stjupɪd/, or the Spanish word *muerto*, /mwerto/, for instance, we could avoid taking a stand on whether the /j/ or /w/ is in the onset or the nucleus of the first syllable. Since there is evidence for *both* possibilities, this fence-sitting could be more reasonable than taking one side or the other. Similarly, a finite-state automaton does not require us to decide whether an intervocalic consonant is in one syllable or the next (and might easily allow both possibilities: in Figure 18.1, intervocalic /p t k l ɾ/ or /n/ are acceptable after states 6 (in the Coda 1 block) and 9 (in the Onset 2 block), implying two possible analyses, for example, *pla.ta* vs. *plat.a*. If it is required to eliminate ambiguity, it will be necessary to add a further mechanism to choose between them (see below), but the "generate and test" strategy of constraint-based approaches require such multiple possibilities to be enumerated in order to choose between them.

A second approach to suprasegmental structure in finite-state automata is to employ extra tapes or tiers to label the prosodic status of vowels and consonants, as in Figure 18.4. This is illustrated by the approach taken by Ellison in example 10, below.

Figure 18.4

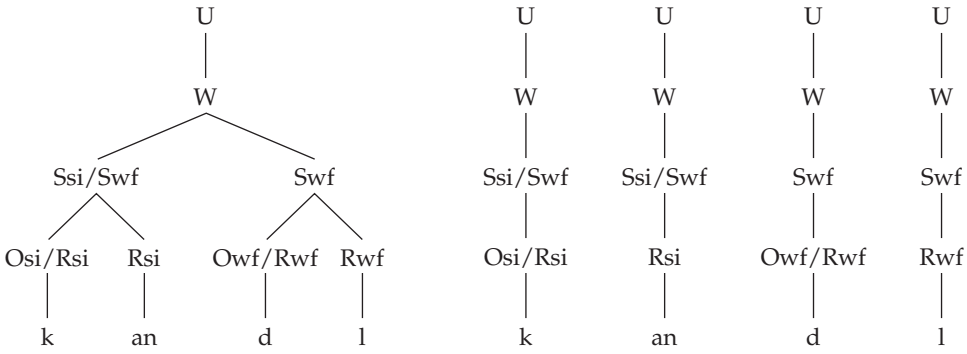


Figure 18.5 A prosodic tree (left) split into its component root-to-frontier paths (right).

Coleman and Pierrehumbert (1997) proposed a version of this method: prosodic structure trees (such as that in Figure 18.5, left) can be divided into vertical root-to-frontier paths (Figure 18.5, right) and then treated as single, complex symbols, for example quadruples such as (W, Swf, Owf/Rwf, d).

In Figure 18.4 and 5, prosodic trees have a fixed number of hierarchical levels. A third method for representing and processing prosodic structure using finite-state automata without multiple tapes is to employ brackets, as in the "bracketted grid" approach to metrical structure (Halle and Vergnaud 1987). Idsardi (2009) shows how his proposed rules for computing metrical structure may be implemented using finite-state transducers; for a slideshow that demonstrates his approach in a more dynamic fashion, see Idsardi (2004).

Speech technology makes extensive use of finite-state transducers. If instead of *alphabetic* phonetic symbols we use acoustic representations of slices of the speech signal, such as spectral vectors, we can relate speech signals to their phonemic transcriptions. Because of the large number of distinct states of a signal, the construction of a transducer that relates acoustic features to segmental labels has to be automated. Hidden Markov Models (HMM's), in which a probability distribution is associated with each state transition, are an important extension of this approach. The probability distributions are determined empirically, by training the automaton on pairings of known signals with their transcription (see Coleman 2005: 144–149). Such devices are usually used for automatic speech recognition (Rabiner and Juang 1993), though HMM's can also be used to generate a signal from a transcription (Donovan 1996). Some attempts have been made to use multi-tiered, autosegmental representations, encoded as synchronized finite-state automata, for automatic speech recognition, with some success (e.g. Sun and Deng

2002). The addition of probabilities to state transitions provides a means for resolving ambiguity in cases where there is more than one possible path through an automaton: using the Viterbi algorithm, find the state sequence with the highest overall probability (Coleman 2005: 207–208).

3 Constraint-based Approaches to Phonology

Finite-state techniques were largely developed outside the arena of *non*-computational phonological research. Another alternative to transformational grammars began to be explored in the 1980s: constraint-based phonology. This term, which embraces diverse variants, is now widespread. In constraint-based approaches such as Optimality Theory (the best-known variant), lexical entries underspecify the surface forms of words: that underspecification permits a certain range of variation. For example, if stresses are not stored in the lexicon, a language may permit a range of variant stress patterns in different contexts, for example, *tórrĕnt* vs. *tŏrrént(-ĭăl)*, *tŏrrènt(-iálity)*. Likewise, if syllable structure is not stored, a word might be syllabified in different ways in different contexts, for example, *feel* (syllable-final /l/) vs. *fee.ling* (syllable-initial /l/). Which stress pattern or syllabification occurs in which context is determined by a set of interacting constraints. Taking Optimality Theory as a concrete example, the derivation of a surface form from the lexicon proceeds as in Figure 18.6.

This approach to computation is called "generate and test": starting from a single lexical entry, generate numerous extensions or alterations (the "candidates") and then select the one (or the subset) that is best, according to a set of constraints.[9]

In constraint-based approaches, the specification of languages is kept apart from questions of how computations such as generation, recognition, and translation are performed. Linguistic properties can be specified by declarative constraints, such as "a word consists of one or more syllables," "a syllable consists of an onset,

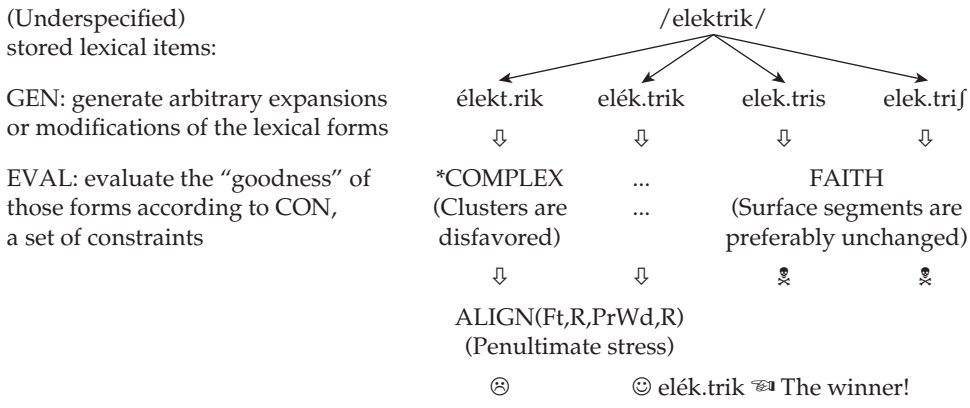


Figure 18.6

a nucleus, and a coda," "/b/ can be a coda," "/a/ is a nucleus," or prohibitions such as "/b/ is not a nucleus." Such constraints define a set of "words," such as /bab/, /babbab/, /babbabbab/ and so on, and a set of nonwords, such as /bbbb/. If constraints are expressed as propositions of logic, questions such as "Is /babbabab/ a word?" and "What is the set of words that end with /k/?" may be determined using *automated deduction*, computational methods for answering questions (or proving facts) based on a set of facts or declarations, that is, implementations of logical deduction, as opposed to "first do this, then do that" rule-following. The best-known example of this approach to computation is the declarative programming language Prolog (Clocksin and Mellish 1981), the release of which had a profound influence on declarative grammar in the 1980s.

Most phonological rules are easily expressed as constraints. For example, a *feature-filling* rule such as [+nasal] → [+voice] can be interpreted as "if *x* is [+nasal] then *x* is also [+voice]." Because of the similarities between the deductive approach to computation and current thinking in phonology, constraint-based methods are a popular approach to "pure" computational phonology, in that they implement non-computational phonological frameworks relatively faithfully, rather than introducing new formalisms (such as finite-state automata or probabilities, etc.) that are not generally employed by non-computational phonologists.

3.1 Declarative Phonology

I mentioned above that one path away from the problems of transformational phonology emulated Gazdar *et al.*'s (1985) strategy of eliminating transformations and the distinction between underlying and surface representations. GPSG provided a formalism for feature structures, feature cooccurrence restrictions, feature specification defaults, and a seamless combination of hierarchical constituent structure and features. It showed that feature cooccurrence restrictions need not be limited to the [A] → [B] format of Chomsky and Halle but, being logical formulae rather than rewriting rules, might just as well be biconditional [A] ↔ [B] or include negations (e.g. [BAR 1] ⊃ ~[SUBCAT]), offering us a model on which to write phonological constraints such as ~[+nasal, –voice]. (Any procedural connotations of an implication such as [+nasal] → [+voice] can be immediately eliminated by noting its logical equivalence to ~([+nasal] & ~[+voice]), which simplifies to ~[+nasal, –voice].)

The value of feature *structures* (as opposed to the unstructured *lists* of features of SPE segmental phonology) had been informally explored in the 1970s by, for example, Lass (1976: 154–155), and a hierarchical structure of functional groups of features proposed by Clements (1985) (Figure 18.7) was rapidly taken up in the field. GPSG-inspired phonologists quickly saw that such "feature geometry" could easily be formalized using feature structures; for example, Figure 18.8, from Klein (1987).

In declarative approaches to grammar, a distinction is drawn between *descriptions* of structures and the structures themselves, most linguistic diagrams being understood as descriptions. This means that phonological alternations can also

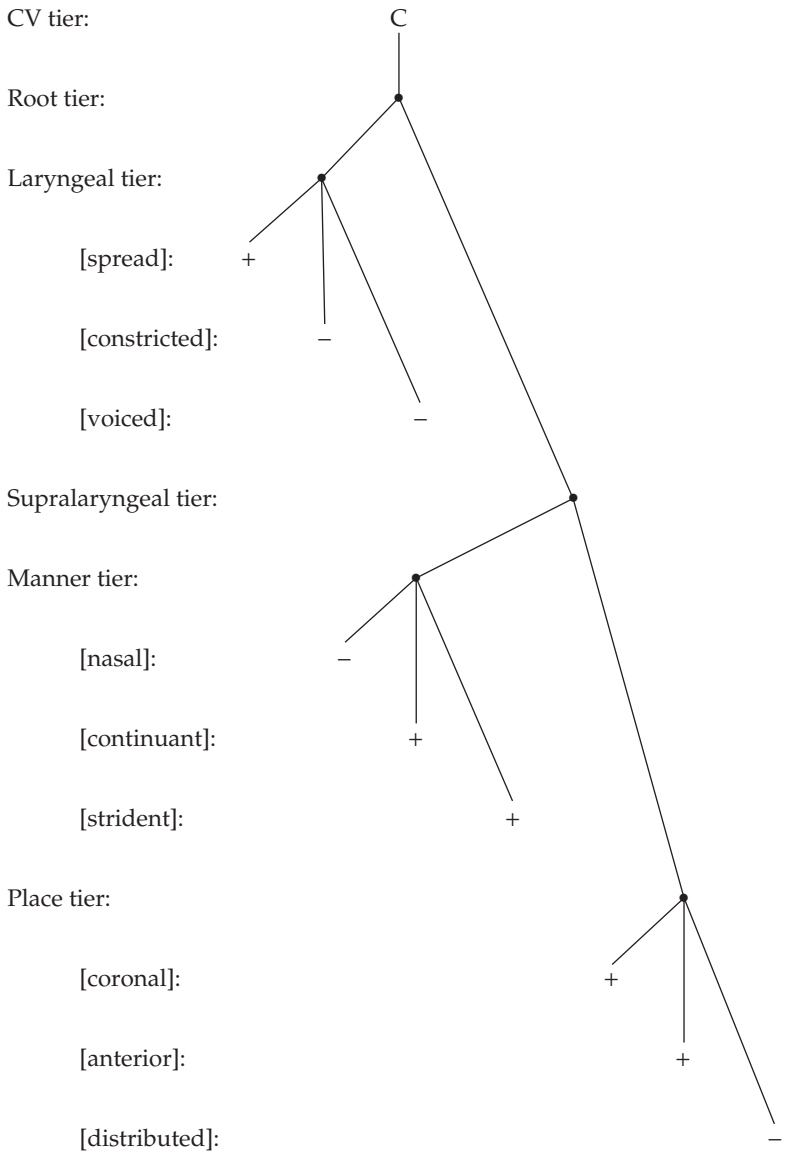


Figure 18.7 Hierarchical structure of phonological features of [s] (Clements 1985).

be expressed using feature structures, as in Figure 18.9, from Calder (1988), who at the time was one of Klein's doctoral students.

Research into declarative phonology, morphology, and syntax in the 1980s was underpinned in the UK and Europe by increased funding for computational linguistics, part of the international response to the Japanese Ministry of International

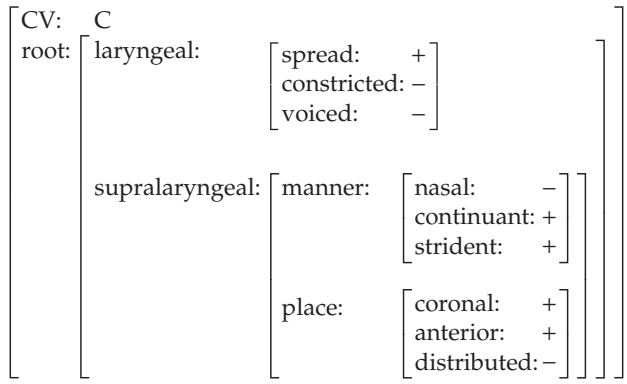


Figure 18.8 Feature structure encoding of Clement's feature tree for [s] (Klein 1987). Reprinted with permission from Ewan Klein.

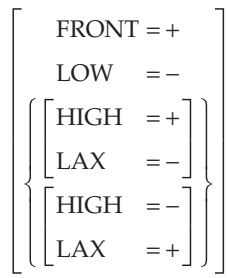


Figure 18.9 Feature-structure description of alternating /i/ ~ /ɛ/ in "economic" (Calder 1988).

Trade and Industry's "Fifth Generation" consortium. In the UK at this time syntactic and phonological theory were hardly separate from computational linguistic research, as many of the same people were doing both. Klein and Calder had funding from the European ESPRIT programme; some other GPSG researchers received support under the UK government's Alvey programme. New start-ups and private corporations big enough to support research and development in computational linguistics also sustained declarative grammar development, for example, SRI, Xerox, British Telecom, and AT&T.

Work on feature structures and their potential applications to phonology was accompanied by work on finite-state models in particular. One factor driving this was that GPSG syntactic theory offered nothing akin to the *non*-hierarchical data structures needed to implement autosegmental phonology.

Although hierarchical structures such as syllable structures or metrical trees of limited depth can be processed using finite-state techniques, they are also easily defined using context-free phrase structure grammars, and then processed using tools such as standard parsing algorithms for context-free languages (Church

1983; Randolph 1989; Coleman 1992). Top-down parsing, in particular, provides an algorithmic implementation of the template-driven approach to syllabification advocated by theoreticians such as Selkirk (1982). To deal with our Spanish examples, for example, the context-free grammar (9) is reasonable.

(9)

$W \rightarrow \sigma_0$	$C_1 \rightarrow$ p | t | k | b | d | g | f | θ | s | x | ʧ | ɾ | l | ʎ | m | n | ɲ
$W \rightarrow \sigma_1\ \sigma_2$	$C_2 \rightarrow$ p | t | k | b | d | g | f
$\sigma_0 \rightarrow O\ R_0$	$C_3 \rightarrow$ p | k | b | g | f
$\sigma_1 \rightarrow O\ R_1$	
$\sigma_2 \rightarrow O\ R_0$	
$\sigma_2 \rightarrow O\ R_1$	$R_1 \rightarrow N$
$O \rightarrow \emptyset$	$R_1 \rightarrow N\ Co_1$
$O \rightarrow C_1$	$R_1 \rightarrow N\ Co_2$
$O \rightarrow$ i | u	$R_1 \rightarrow N\ Co_3$
$O \rightarrow C_1$ i | u	$R_1 \rightarrow N\ Co_1\ Co_3$
$O \rightarrow C_2$ ɾ	$N \rightarrow$ i | e | a | o | u
$O \rightarrow C_3$ l	$Co_1 \rightarrow$ p | t | k | i | u | l | ɾ | n
$R_0 \rightarrow N\ Co_1$	$Co_2 \rightarrow$ m
$R_0 \rightarrow N\ Co_2$	$Co_3 \rightarrow$ s

According to this grammar, our three well-formed test words, *buey*, *tiempo*, and *plata* can be parsed as in Figure 18.10 (upper part). The initial substrings *ml . . .* and *sl . . .* cannot be parsed by any combination of the rules, however, so all possible derivations of, for example, *mleta* or *sleta* crash, and we conclude that they are ill-formed.

As with finite-state approaches, context-free grammars of syllable structure are typically ambiguous, defining multiple possible parses for most strings, for example, *tiem.po* vs. *tiemp.o*. The standard proposal to resolve this, the Maximal Onset Principle, is usually described in procedural terms ("first, parse as much as possible into onset positions"), but this is not necessary: it is equally effective to add a declarative well-formedness constraint prohibiting the combination of a filled coda and empty onset, as in Coleman (1996: 204). Preventing ambiguity by the addition of *ad hoc* constraints is inelegant, however, and may be empirically undesirable: for example, though /kr/ is a perfectly good onset in English, the glottaling of /k/ in for example, the pronunciation of the surname *Ackroid* by some speakers indicates coda /k/+onset /r/, whereas the aspiration of /k/ and devoicing of /r/ in *Ukraine* indicates onset /kr/. To exploit the ambiguity of a phonotactic grammar, an effective mechanism is needed to select the *preferred* parse for any given case. One mechanism for achieving this is to use *probabilistic* grammars (see e.g. Coleman 2000; Müller 2002) to determine the most likely analysis. Thus, although the probabilistic grammar in Coleman (2000) maximizes *most* clusters as onsets, for example *U.kraine*, it assigns higher probability to, for example, *Ack.roid*, *mush.room*, and *Cots.wold*, contrary to the uniformity predicted by the Maximal Onset Principle.

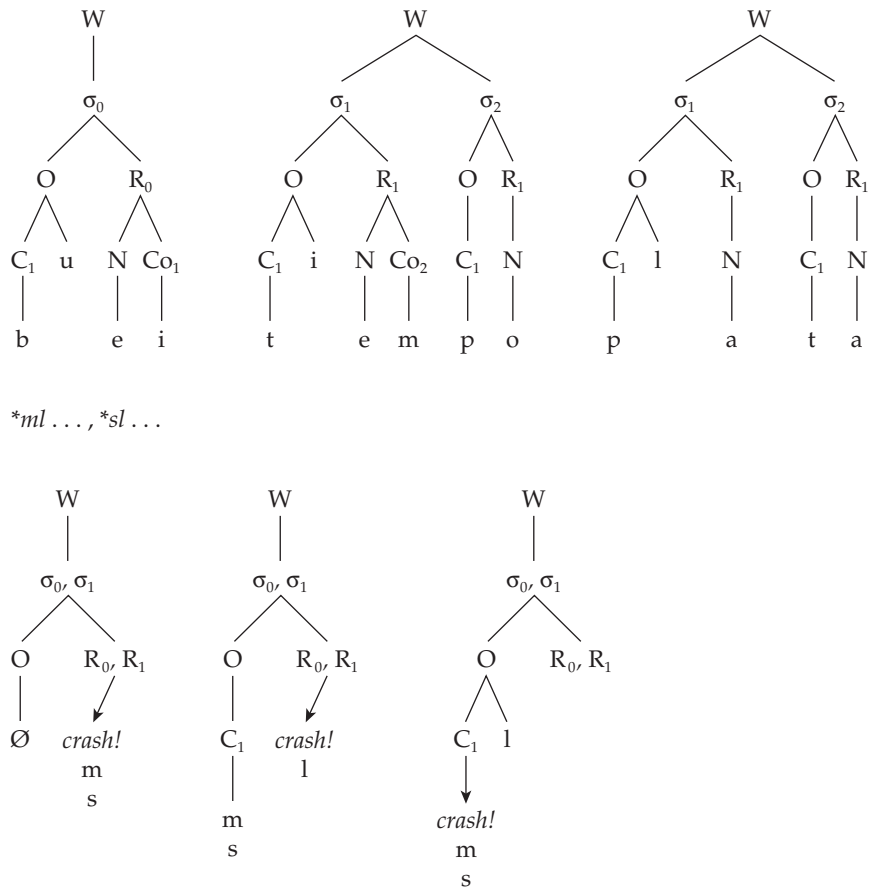


Figure 18.10 Above: parse trees defined by a context-free grammar of Spanish phonotactics. Below: derivations of **mleta*, **sleta* cannot be completed (crash), showing that those items are ill formed.

3.2 Optimality Theory

Optimality Theory offers another, usually non-probabilistic, method for ranking the possible parses. In many constraint-based formalisms, constraints must be consistent, without contradictions. However, Optimality Theory permits conflicts between constraints, which express defaults or tendencies rather than exceptionless regularities. A partial ordering over the set of constraints defines a priority ranking of the constraints.

To illustrate the application of Optimality Theory to our benchmark Spanish examples, we follow the analysis of Shepherd (2003),[10] augmented by details from the foundational Optimality Theory literature (e.g. Prince and Smolensky 2004 [1993]). First, consider the analysis of *buey* (/buei/), which in fact consists of a

single syllable, but which in theory might be parsed in a large number of different ways, yielding up to four syllables: buei, b.uei, bu.ei, bue.i, b.u.ei, b.ue.i, bu.e.i, or b.u.e.i. Although Optimality Theory phonologists have been extremely unforthcoming about the details of the Gen function, all these candidates can easily be generated from the lexicon by inserting syllable boundaries between any pair of segments. The second challenge is to determine whether the lexical segments /u/ and /i/ are to be parsed as nuclear or non-nuclear constituents. The latter is the desired outcome, so that /u/ is to be pronounced as non-nuclear [w] and /i/ as non-nuclear [j].

The output of Gen must provide not only information about syllable divisions, but also about the affiliation of segments to syllable constituents. Following Prince and Smolensky (2004 [1993]), we mark syllable nuclei with an acute accent. Segments preceding the nucleus are assumed to be in the onset and those following the nucleus are taken to be in the coda; thus, pre-nuclear /u/ will be [w] and post-nuclear /i/ will be [j]. Where the first segment of a syllable is a nucleus, an empty onset □ is hypothesized, and where the syllable-final segment is the nucleus, an empty coda □ is hypothesized. Gen also allows segments to be left unparsed, a possibility that we shall ignore in the examples that follow. For the monosyllabic candidate /buey/, therefore, we just consider four possible syllable structures: /□b́uei, búei, buéi, bueí□/.

Shepherd's approach to Spanish syllabification, like much earlier work, is driven by a combination of structural constraints and constraints on sonority (sonority sequencing and minimum sonority distance in clusters). The main relevant constraints are the following:

MSD-2$^{\text{Ons}}$:	The minimum sonority distance between the two elements of a complex onset is 2.	
SonSeq:	Onsets must rise in sonority towards the nucleus and codas must fall in sonority from the nucleus.	
Onset:	*[$_{\sigma}$V	"Syllables must have onsets."
No-Coda:	*C]$_{\sigma}$	"Syllables must not have codas."
*Complex$^{\text{Ons}}$:	*[$_{\sigma}$CC	"Onsets are simple."
*Complex$^{\text{Cod}}$:	*CC]$_{\sigma}$	"Codas are simple."[11]

For Spanish, these are ranked by Shepherd into the hierarchy Onset, MSD-2$^{\text{Ons}}$, SonSeq >> *Complex$^{\text{Cod}}$, No-Coda >> *Complex$^{\text{Ons}}$. Table 18.3 shows how these constraints evaluate the four monosyllabic candidate parses.

In brief, given monosyllabic parses, (a), (b) and (d) fall at the first hurdle because they violate sonority sequencing (and, furthermore, b́uei lacks an onset[12]), so (c), buéi is the optimal monosyllabic parse. The other possible syllabification candidates are assessed in Table 18.4. (We need not persist in considering parses which violate sonority sequencing, that is, those with b́u, -úe, úe or eí, so SonSeq and MSD-2$^{\text{Ons}}$ are irrelevant to the remainder of the candidates.) With Onset and *Complex$^{\text{Ons}}$ doing the key work, b́□.uéi comes out as more optimal than the desired monosyllabic parse, buéi. Clearly, something is not right: a constraint against

Table 18.3 Evaluating monosyllabic candidate parses of *buey*.

/buei/	Onset	MSD-2^{Ons}	SonSeq	*Complex^{Cod}	No-Coda	*Complex^{Ons}
a. □b́uei	*		**	**	*	
b. búei			*!	*	*	
c. buéi					*	*
d. bueí□			*!			**

Table 18.4 Evaluating polysyllabic parses of *buei*.

/buei/	Onset	*Complex^{Cod}	No-Coda	*Complex^{Ons}
c. buéi			*	*!
e. ☞ bə́.uéi			*	
f. bú.éi	*!		*	
g. bué.í	*!			*
h. bə́.ú.éi	**!		*	
i. bə́.ué.í	**!			
j. bú.é.í	**!			
k. bə́.ú.é.í	***!			

empty nuclei is needed that must dominate at least *Complex^{Ons}. In general, parses with fewer syllables are better as they tend to have fewer violations of Onset, and monosyllabic buéi is best of all because it has a good sonority profile.

Next, consider the syllabification of *tiempo*, which should be disyllabic. In Table 18.5, all monosyllabic parses are enumerated in (a) to (f), some of them quite curious ones with syllabic obstruents or coda /o/. Every disyllabic partition of the string is enumerated in (g) to (m) and every trisyllabic partition in (n) to (w). Interestingly, although the correct analysis, (l), is optimal, this depends solely on the *unranked* pair of constraints Onset and SonSeq: the other four are irrelevant. (MSD-2^{Ons} is irrelevant in this case because every form that violates it also violates Onset or SonSeq or both.)

3.3 Computing Optimality Theory

Ellison (1994) showed how Optimality Theory derivations can be computed using finite-state transducers. First, lexical items can be translated into regular expressions

Table 18.5

/tiempo/	Onset	MSD-2^{Ons}	SonSeq	*Complex^{Cod}	No-Coda	*Complex^{Ons}
a. ◌́tiempo	*		***	*	*	
b. tíempo			**	*	*	
c. tiémpo			*	*	*	*
d. tiḿpo		*	**	*	*	*
e. tiémṕo		**	***		*	*
f. tiempó		**	**			*
g. ◌́t.◌íempo	**		**	*	*	
h. ◌́t.iémpo	*		*	*	*	
i. tí.◌émpo	*		*	*	*	
j. tíe.mpó		*	**		*	*
k. tié.mpó		*	*			**
l. ☞ tiém.pó					*	*
m. tiémp.◌ó	*			*	*	*
n. ◌́t.◌í.◌émpo	***		*	*	*	
o. ◌́t.ié.mpó	*	*	*			*
p. ◌́t.iém.pó	*				*	
q. ◌́t.iémp.◌ó	**			*	*	
r. tí.◌é.mpó	*	*	*			*
s. tí.◌ém.pó	*				*	
t. tí.◌émp.◌ó	**			*	*	
u. tié.◌ḿ.pó	*					*
v. tié.◌ḿp.◌ó	**					*
w. tiém.◌ṕ.◌ó	**				*	*

encoding the range of possible variant outputs of Gen. For example, 64 different possible syllabifications of the Arabic segmental sequence *alqalamu* can be expressed as the regular expression in (10).

(10) $\left\{\left|\begin{matrix}O\\0\end{matrix}\right|\right\}$ $\begin{matrix}NCON\\a\ l\ q\ a\end{matrix}$ $\left\{\left\{\left|\begin{matrix}C\\0\end{matrix}\right|\right\}\left|\begin{matrix}O\\l\end{matrix}\right|\begin{matrix}C\\l\end{matrix}\left\{\left|\begin{matrix}O\\0\end{matrix}\right|\right\}\right\}\begin{matrix}N\\a\end{matrix}\left\{\left\{\left|\begin{matrix}C\\0\end{matrix}\right|\right\}\left|\begin{matrix}O\\m\end{matrix}\right|\begin{matrix}C\\m\end{matrix}\left\{\left|\begin{matrix}O\\0\end{matrix}\right|\right\}\right\}\begin{matrix}N\\u\end{matrix}\left\{\left|\begin{matrix}C\\0\end{matrix}\right|\right\}\right\}$

The brackets group alternative possibilities, separated by vertical bars. To explain this further, note that

$$\left\{\begin{matrix} \mathrm{C} \\ 0 \end{matrix}\middle|\right\}$$

means "an empty coda, or nothing" (there is nothing to the right of the bar), which is equivalent to saying "the coda must be empty." Therefore,

$$\left\{\left\{\begin{matrix} \mathrm{C} \\ 0 \end{matrix}\middle|\right\}\begin{matrix} \mathrm{O} \\ \mathrm{l} \end{matrix}\middle|\begin{matrix} \mathrm{C} \\ \mathrm{l} \end{matrix}\left\{\begin{matrix} \mathrm{O} \\ 0 \end{matrix}\middle|\right\}\right\}$$

means "an empty coda followed by an onset l, or a coda l followed by an empty onset." In this way, the two possible syllabifications of the intervocalic l are spelled out. Using this notation, the 23 possible syllabifications of *tiempo* of Table 18.3 can be packed up as in (11).

(11) $$\left\{\left\{\begin{matrix} \mathrm{O} \\ 0 \end{matrix}\middle|\right\}\begin{matrix} \mathrm{N} \\ \mathrm{t} \end{matrix}\middle|\begin{matrix} \mathrm{O} \\ \mathrm{t} \end{matrix}\left\{\left\{\left\{\begin{matrix} \mathrm{O} \\ 0 \end{matrix}\middle|\right\}\begin{matrix} \mathrm{N} \\ \mathrm{i} \end{matrix}\middle|\begin{matrix} \mathrm{O} \\ \mathrm{i} \end{matrix}\right\}\middle|\begin{matrix} \mathrm{C} \\ \mathrm{i} \end{matrix}\left\{\left\{\left\{\begin{matrix} \mathrm{O} \\ 0 \end{matrix}\middle|\right\}\begin{matrix} \mathrm{N} \\ \mathrm{e} \end{matrix}\middle|\begin{matrix} \mathrm{O} \\ \mathrm{e} \end{matrix}\right\}\middle|\begin{matrix} \mathrm{C} \\ \mathrm{e} \end{matrix}\left\{\left\{\left\{\begin{matrix} \mathrm{O} \\ 0 \end{matrix}\middle|\right\}\begin{matrix} \mathrm{N} \\ \mathrm{m} \end{matrix}\right| \ldots$$

Each constraint is modeled as a finite-state transducer taking as input such regular expressions and mapping them to a list of integers, the order of which corresponds to the sequence of constraints; each integer encodes how many times the input violates the respective constraint. Thus, a candidate with the list (–2, –1) violates the first constraint twice and the second once. Such lists can be compared in order to assess the comparative well-formedness of two or more particular candidate forms; for example, (–10, –31, –50) is better than (–10, –34, –12) because –10 = –10, but –31 > –34. Karttunen (1998) describes a specific computational implementation in this vein (and Tesar 1996 a variant approach), and Frank and Satta (1998) provide a formalism.

As well as occupying (possibly empty) syllable positions, a segment may be free (unparsed), as in the case of extrametrical or extrasyllabic material. Since every segment must either be parsed as O, N, or C, or unparsed, we can represent all possible parses of any string using a finite-state automaton of the general structure of Figure 18.11.

In Figure 18.11, each state (circle) represents the possibility that a particular input segment is parsed as an O, N, C, or unparsed; that is, the columns represent the four separate possibilities at each step, and the rows represent the succession of input segments. Clearly, for each input segment there are always 16 possible transitions to the machine's state when the next symbol is input. Also there are four ways of dealing with the first symbol, which we can consider as four transitions (not shown here) from the initial boundary to the first input symbol. Likewise, there are four ways of moving from the last input symbol to the end of the string.[13] Some of the possible transitions lead to constraint violations.

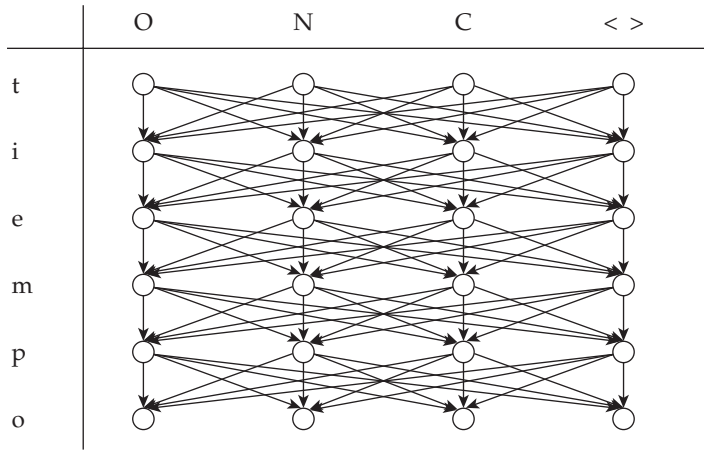


Figure 18.11

Optimality Theory is founded upon the claims that a unique set of universal constraints applies in all languages, and that differences between one language and another are due to different rankings of the constraints. For the child learning a language, therefore, the task is to discover the appropriate ranking as well as the underlying forms. These problems have also been examined and modeled computationally. A method for learning the ranking of constraints – the Constraint Demotion Algorithm – is presented by Tesar and Smolensky (1998), Tesar (1998) and in other work by those authors. Boersma and Hayes (2001) offer an alternative learning algorithm embodying probabilistic variation in the rank order of each constraint, the Gradual Learning Algorithm. A comparative evaluation of the two approaches is given by Keller and Asudeh (2002). For learning underlying forms (or for finding underlying forms given a surface form and a constraint ranking), see Riggle (2004). Since the debate on the merits of these various proposals is still current, I merely guide the reader to that growing literature.

Some aspects of Optimality Theory may at first seem difficult to swallow: for example, possibly infinite sets of candidate forms frequently raises objections from critics. But when the space of solutions is structured and searched in a sensible way, the fact that there are infinitely many possible candidates does not require us to enumerate and evaluate them all – which is clearly impossible. It is sufficient to be able to consider the best *n* candidates, for finite *n*, from which to select the best one.

In all constraint-based approaches, the preferred analysis is the most "harmonic" analysis, the one that fits together the best (Smolensky 1986; Goldsmith 1993a). Declarative phonology claims that the optimal, most harmonic analysis is one in which *no* constraints *ever* conflict; it achieves this by allowing language-specific constraints, whereas Optimality Theory values universality of constraints above "harmony." Simultaneous satisfaction of multiple, interacting, parallel constraints – "harmony" – is also found in connectionist approaches.

4 Connectionist Approaches

Connectionist models arose from studies of the computational abilities of neural networks, an attempt to understand the workings of neural circuitry.[14] Although research into artificial neural networks was pursued throughout the twentieth century, interest in connectionist models developed greatly in the 1980s, when it became easy to implement them in software rather than as electronic circuitry.[15] By that time, computational modeling of language and other aspects of cognition (Artificial Intelligence, or "A.I.") was quite advanced (see e.g. Winograd 1983 or Allen *et al.* 1987 for examples of the then state-of-the-art). But for all their sophistication, such models had severe problems: they had to be painstakingly programmed – they did not learn – and they were brittle: they did not fare well with input lying outside the range they had been programmed to deal with. In contrast, connectionist models require relatively little programming, learn their behaviors on the basis of training from example data, and can work reasonably well on degraded or novel input. Even to its critics, these strengths made connectionism a rather impressive challenge to the more mature methods of rule-based A.I.

Connectionist models are well-suited to the computation of relationships between distinct levels of representation, especially when the correspondence between the two levels is unclear or ill-formalized, because (like HMM's) they can *learn* correspondences between representations by presenting them with numerous examples of the relation in question. This makes it unnecessary to discover and debug a list of phonological rules, a task which often yields unforeseen errors. Faced with the difficulty of finding a complete and correct set of rules, such as, phoneme to allophone translation, *training* a model is an appealing prospect. Connectionist models have been employed with some success for grapheme to phoneme translation (Sejnowski and Rosenberg 1987), recognition of phonemes from speech acoustics (Waibel *et al*. 1989), learning syllabification and stress patterns (Larson 1992; Daelemans and van den Bosch 1992; Gupta and Touretzky 1994), predicting the next phoneme in a string of phonemes (Elman 1990), and acquiring phonological representations from semantic, acoustic, and articulatory representations (Plaut and Kello 1999).

The most influential works in connectionism were the two-volume collection of papers on parallel distributed processing by Rumelhart *et al.* (1986) and McClelland *et al.* (1986); the TRACE model of speech perception (McClelland and Elman 1986), which combined acoustic speech input with word recognition; and the model proposed by Rumelhart and McClelland (1987) for learning the (regular and irregular) past tenses of English verbs. The fact that this work drew lengthy critical responses from such defenders of the symbol-and-rule-based approach as Fodor and Pylyshyn (1988) and Pinker and Prince (1988) testifies to the seriousness with which the connectionist challenge was received. But in spite of Prince's rejection of connectionism *per se*, some aspects of it stuck: the idea of parallel, interacting constraints relating co-present levels of representation (rather than an earlier input leading to a later output); Smolensky's presentation of harmony – the optimal but not necessarily perfect satisfaction of parallel constraints – and the retention

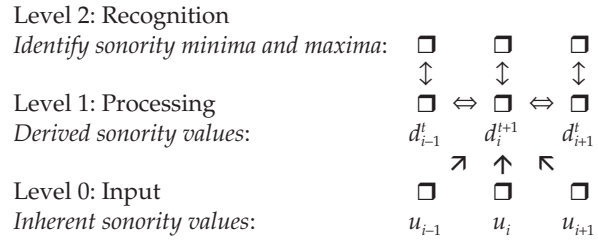


Figure 18.12 Larson's syllable network. Reprinted with permission from Gary Larson.

of purely symbolic representations and constraints – a brew which produced Optimality Theory. Although most linguists paid little interest to connectionism, note that many psychologists of language were impressed by it (e.g. Dell 1986; Gaskell and Marslen-Wilson 1996; Plaut and Kello 1999).

Connectionist treatments of syllabification and phonotactics in Spanish are presented in Larson (1992). Larson's syllable network employs three ranks or levels of units, with some excitatory/inhibitory connections between neighboring units and/or units on different levels (Figure 18.12).

At the input level of the network, each unit corresponds to one segment, including leading and trailing boundary symbols, #. Thus, processing the six-segment word *tiempo* requires a network eight units wide. The approach to syllabification employed here is based on sonority: the activation level of each unit at level 0 represents the *inherent* sonority of each input segment, from which the network computes at level 1 the *relative* sonority profile of a word, from which maxima (syllable peaks) and minima (syllable edges) can be picked out at level 2. The computation can be considered successful or correct if (a) the number of syllable peaks is correct, and (b) there are no spurious syllable peaks in places where native speakers would not place them. Syllable divisions can be derived from derived sonority minima, as we shall see.

In Larson's early experiments, inherent sonority values at the input level were simply read from a table embodying a "universal sonority hierarchy," in which voiceless stops have inherent sonority $u = 1$ and low vowels have inherent sonority $u = 9$. (In Larson (1992: Chapter 6), sonority values are learned from a corpus.) For #*tiempo*#, therefore, the input vector u could be[16] (0, 1, 7, 8, 5, 1, 8, 0). If syllabifications are simply "read off" such inherent sonority profiles, however, they will frequently be wrong, with sonority peaks where there is no syllable nucleus (as in the initial *f* of words beginning with *fl-* or *fr-*), or with no differentiation between peak high vowels (as in *cinco*) vs. non-peak onglides (as in *tiempo*, in which *i* is [j]). *Derived* sonority, computed at level 1, rectifies such problems, giving an *adjusted* profile in which the sonority of each element can be raised or lowered by some proportion of the sonority of its immediate neighbors, according to the time-dependent equation:

(12) $d_i^{t+1} = u_i + \alpha \cdot d_{i+1}^{t} + \beta \cdot d_{i-1}^{t}$

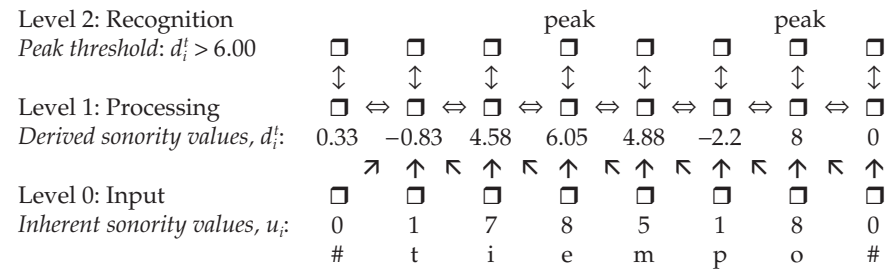


Figure 18.13 Result of syllabification network computations on Spanish input *tiempo*. Reprinted with permission from Gary Larson.

repeatedly recalculated for each unit at time steps from $t = 1$ to n, until the change from d_i^t to d_i^{t+1} is so small that it falls below some small, arbitrary level Δ for all level 1 nodes.

The end-point of this network's computations on the Spanish input *tiempo* is shown in Figure 18.13. With coefficients $\alpha = -0.4$, $\beta = 0$ and a syllable peak threshold $d_i^t > 6.00$, the network correctly determines that (a) the word has two syllables, not three, that is, (b) *i* is not a peak, but (c) is an on-glide to the first syllable peak, *e*.

Similarly, in Spanish *autor*, the *u* is an off-glide, not a peak, that is, *autor* is disyllabic, and in *buey* (with *y* given the same inherent sonority as *i* at the input level), *u* is an on-glide and *y* an off-glide, that is, *buey* is monosyllabic.

In contrast to boundary- and constituency-based approaches to syllabification, the sonority-based approach places syllable boundaries (sonority minima) *on* a particular segment, such as the *p* of *tiempo*, or the *t* of *autor*. While it is possible to interpret the sonority trough as the *first* segment of a syllable (i.e. place a boundary before it), this remains an arbitrary stipulation – as arbitrary as any of the rules proposed in other approaches, and as arbitrary as declaring that segments at sonority troughs are ambisyllabic or syllable-final. Furthermore, the sonority profile does not provide a parse of the syllables into sub-syllabic constituents. As with finite-state automata, sonority does not tell us whether the glide *i* in *tiempo* is part of a rising diphthong in the nucleus, or part of the onset.

4.1 Phonotactics

The phonotactic ill-formedness of **mleta* and **sleta* in Spanish presents a difficulty for any account based on sonority: since *m* and *s* are less sonorous than *l*, *ml-* and *sl-* should be acceptable onsets. Again, the dynamic computation of *derived* sonority is called into play. The small negative value of the α coefficient (−0.4) has an inhibitory effect, lowering the derived sonority of *l* from its high initial value, until the values in Table 18.6 are reached.

Since the peak values of *m* and *s* before *l* are below the threshold of 6, they can be picked out as "false" peaks from which the ill-formedness of **mleta* and **sleta* follows. For *plata*, however, the derived sonority of *p* remains below that of the following *l*, hence *pl-* is a well-formed Spanish onset.

Table 18.6

**mleta*		#	m	l	e	t	a	#
	Input:	0	5	6	8	1	9	0
	Derived:	–1.44	**3.2**	1.78	9.37	–3.05	9	0
**sleta*		#	s	l	e	t	a	#
	Input:	0	3	6	8	1	9	0
	Derived:	–0.99	**2.2**	1.78	9.37	–3.05	9	0
cf. *plata*		#	p	l	a	t	a	#
	Input:	0	1	6	9	1	9	0
	Derived:	–0.08	0.21	1.98	10.04	–2.6	9	0

Since initial clusters /ml/ and /sl/ are well-formed in other languages, and yet the input sonority values in this model are proposed as universal, the language-specific prohibition against these clusters in Spanish depends in part on the threshold value set for syllable peaks. For example, if the threshold is lowered to 3, the network discussed above would continue to accept *plata* as disyllabic and to reject **sleta*, but would now accept *mleta* as trisyllabic *m.le.ta.* In order to accept /ml/ or /sl/ as onsets, however, it would be necessary to alter the value of α (and/or β) so that the derived sonority of /l/ is not so low.

In speech technology, connectionist approaches do not perform as well in speech recognition as HMMs[17] (Trentin and Gori 2001). However, from a cognitive science perspective, the similarities between the performance of these models and certain particular aspects of human behavior is most interesting, as are the parallels in the design and operating principles of artificial and natural neural networks.

5 Reflections

Throughout the preceding half century, phonologists' thinking about the nature of phonological computations (whether in abstract generative grammars or in more concrete models of language processing) has been well informed by the computational techniques of the day. If we step back from particulars and attempt to take a broad view of the subject, and by looking at today's emerging trends, we can discern some of the directions in which computationally-informed phonology is now headed.

Optimality Theory, connectionist phonology, and declarative phonology are superficially quite different formalisms, but they arose in a common intellectual climate, and share quite a few consensus features (which receive little attention because there is little disagreement about them). First, all three formalisms focus on description of the well-formed and prohibited *surface structures*, the observable *outputs* of a generative phonology (and the given *inputs* to a recognition or parsing system), rather than on "underlying" lexical representations. Second, the *intermediate* levels which were so important to transformational generative grammar have gone. This does not mean that constraint-based phonologies neglect stored,

lexical representations: if anything, the focus on individual levels of a derivation and a unidirectional derivation from lexicon to surface (Optimality Theory remains closest to this model) has given way to an emphasis on *correspondences* between pairs of related forms, such as surface–lexical pairs or surface–surface pairs (like Hooper's "via rules"). Fourth, the number of distinct levels has been debated, but even on this point there is a fair consensus: Optimality Theory, connectionism, and two-level phonology agree that possibly/somewhere distinct surface and lexical representations are needed. Declarative phonologists aimed to work with just a single level of phonological representation, a surface representation of which the lexical representation was a part, but it is not clear that this is sufficient: the "surface" in question proves to be quite abstract. Declarative phonological analyses often exploit either the phonology-phonetics distinction (e.g. Coleman 1992), the description/object distinction (e.g. Calder 1988; see the discussion of Figure 18.9 above), or the default/exception distinction. So, there are a variety of ways in which declarative phonology, too, finds it necessary or convenient to work with more than one level of representation. Fifth, constraints are mostly seen as applying together, embodying the parallel and/or distributed model of computation. Finally, different constraints have different degrees of force, scope, or weight. For example, Optimality Theory constraints are explicitly ranked from "most important not to violate" to "not very important"; links and/or nodes in a connectionist network are assigned different weights; constraints in a probabilistic grammar are assigned different probabilities. Even non-probabilistic declarative grammars have a default vs. exception ranking determined by the subsumption hierarchy, with more specific constraints taking precedence over more general constraints or defaults.

5.1 Storage vs. Processing

Computation is not just processing: it also involves storage. In the 1950s, at the start of generative linguistics, computer storage devices had a small capacity and were so expensive that it seemed impossibly wasteful and frivolous *not* to use rules and instead simply to store multiple related forms of a word. (The IBM 350 Disk File from 1956 could store 5 million characters, insufficient for a large dictionary; magnetic core memory came in blocks of 2 kilobytes and cost around $1/bit.) Halle (1985: 105) remarks:

> one may speculate that space in our memory is at a premium and that we must, therefore, store in our memory as little as possible about the phonetic shape of each word, eliminating as many redundancies as possible and placing maximum reliance on our ability to compute the omitted information.

But how infeasible is it, actually, to simply store multiple forms rather than compute them by rule? It is often observed that in some languages, such as Finnish, word formation is regular and very productive: with a core vocabulary of over 88,000 words in the largest version of the *ispell* spell checking dictionary for Finnish,[18]

an average of 670–680 forms can be derived from each word, taking up to 19 Mb of memory: just a small fraction of, say, a USB memory stick.

Halle and Stevens (1959) considered storage of the many variant *acoustic* forms of each word to be equally if not even more implausible:

> The size of the dictionary in such an analyzer increases very rapidly with the number of admissible outputs, since a given phoneme sequence can give rise to a large number of distinct acoustic outputs. In a device whose capabilities would even remotely approach those of a normal human listener, the size of the dictionary would, therefore, be so large as to rule out this approach.

Today, however, the idea that storage is limited and costly yet robust whereas processing is cheap seems implausible. We store large libraries of music on portable music players: a far greater feat of storage than, say, *uncompressed* storage of one recording of each word in a 65,000-word vocabulary, which will fit on a single 512 MB chip. Contrary to the expectations of the 1950s–1970s, today's automatic speech recognition and speech synthesis devices *do* store huge volumes of acoustic data encoding specific examples of how each word sounds, and (some aspects of) its pronunciation variation. And it is clear that the brain's storage ability is prodigious (though not necessarily accurate): just think, for a moment, of the hundreds of faces and the many hours of music with which you may be familiar. Can you recall the sound of many entire music CD's in your mind? What is the size of your auditory "repertoire"?

Extensive storage of detailed and particular word-forms is now regarded as a credible and interesting hypothesis about spoken language storage. Whereas generative phonologists take it as axiomatic that details of a particular speaker's voice characteristics are *not* stored in the lexicon, Goldinger (1997) and Goldinger and Azuma (2004) present respectable experimental evidence and arguments in support of the proposition that such speaker-specific details *are* in fact stored. In view of its potential to overturn long-entrenched dogmas of theoretical phonology, the growing body of such work on "episodic" or "exemplar-based" lexical memory cannot be ignored, and is inevitably shaping our thinking about phonological computations, motivating a closer look at connectionist and other "content-addressable" or "holographic" models of memory.

5.2 *Corpus-based Work*

Another consequence of the enormous increase in computer storage during the last half century is a resurgence in corpus-based computational linguistics. Grounding in good data has always been considered a virtue, of course; computational linguistics is making it obligatory. With a small, clean dataset from a published grammar or one of the "standard problems" of phonology (e.g. those in Halle and Clements 1983), elegant, general, and complete analyses are possible; such data and analyses are prevalent in the lectures and publications of many professional phonologists and students alike. However, working with corpora, one is

immediately confronted with exceptions, variability due to disparate causes ranging from speech rate to sociolinguistic factors, and the near impossibility of attaining a complete, correct treatment. Indeed "completeness" and "correctness" become ill-defined: one strives instead to produce extensive analyses of a specific part of the data, or to compare two or more different models of some data sample. Inevitably, quantitative methods are necessary: even if the data is symbolic (e.g. transcriptions), items can be counted and relevant statistics computed (cf. Hayes and Londe 2006). Evaluating a model of some data requires a measure of "goodness of fit" to be calculated, and if there is no variation in your data, you probably do not have enough data. For some illustrative examples of corpus-based computational phonology, see, for example, Withgott and Chen (1993) or Patterson and Connine (2001). In short, the use of toy grammars and standard examples is quite inappropriate in a serious science. If this point is accepted, it will be seen that a high proportion of theoretical work in phonology is problematic, one of the reasons for the little regard paid to it in speech science, speech technology, and psychology.

5.3 Technical Literacy

Programmers and computer scientists do not begin each new piece of work from scratch: reusability of methods is a key to good code. Similarly, today's computational phonologist has an armoury of well-tried methods with which to work: the pioneering days of working on a single technique (such as finite-state phonotactics, or context-free syllabification) are perhaps already past. Familiarity with the range of methods presented in handbooks such as Jurafsky and Martin (2000) or McLeod *et al*. (1998) may help tomorrow's phonologists avoid too close an attachment to just one view of phonological computations, and equip them with the technical literacy to build models without too much regard for the fads or fashions of phonological theory. On the advice of Pierrehumbert *et al*. (1996), "as Feynman suggests in his discussion of research in physics, 'we must keep all the theories in our heads.'" The practical ability to implement a phonological analysis as a working piece of software enables the researcher to test it against data, identify its strengths and weaknesses, and quantify its overall performance vis-à-vis competing accounts (see e.g. Coleman 2000; Wedel 2007).

Increased numeracy and technical literacy help break down some of the barriers between old-fashioned *theoretical* phonology and relevant neighboring disciplines: not just computational linguistics, but also experimental phonetics, speech technology, neuroscience, and so on. In turn, it becomes feasible to entertain or even begin to construct more integrated models in which phonological computations are joined up with other aspects of linguistic structure, for example, orthography, prosody, the lexicon, semantics, interaction, acquisition, social factors, or historical change. If my futurology is reasonable, the computational view of phonology in future editions of the handbook will be large-scale, probabilistic, joined up with the rest of language and (to within a certain margin of error) measurably wrong.[19]

NOTES

1 Halle (1975) repudiated this work as a misguided failure: "my elaborate computations of the information content in bits of the different phonemes of Russian . . . have been, as far as I know, of absolutely no use to anyone working on problems in linguistics." But this is rather too pessimistic: Cherry, Halle, and Jakobson (1953) provides a number of interesting statistics on Russian, such as (i) the relative frequency of feature-values, from which the unmarked values [–voiced, –sharp, –stressed, –continuant, –grave] can be inferred; (ii) the number of bits per phoneme – about five or six, demonstrating that their nine distinctive features are technically more than necessary; (iii) the most frequent Russian phoneme is /a/, consistent with Jakobson's universals; (iv) the least frequent phoneme is /g^j/, a new addition to the traditional repertoire of Standard Russian phonemes, noted by Cherry, Halle, and Jakobson (1953: Note 7). In 1975, Halle could not have foreseen the later success of probabilistic approaches in modelling the acoustics-to-phoneme mapping in continuous speech recognition systems, and the consequent statistical revolution in computational linguistics.

2 Chomsky and Halle (1968: 35) employs the notation $\acute{\Sigma}$ as an abbreviation for a stressed syllable, "that is, a string of the form $C_0V^1C_0$."

3 See also Maxwell (1994).

4 The "freightyard diagram" of Hockett (1958: 291), adopted by Systemic Phonologists such as Halliday (1992: 118), is a notational variant of such devices.

5 As I am not a specialist in Spanish phonology, I expect that this automaton will be defective, defining sequences that should not be allowed in Spanish and failing to define others that ought to be included. It is given here just for expository purposes.

6 *Any* strings, including strings in languages that cannot be generated by finite state automata, such as the strictly context-free language a^nb^n. No finite state automaton can correctly *reject* the ungrammatical strings of strictly non-finite-state language, however.

7 This declarative alternative view of rewriting is quite comparable to Optimality Theory's input-output correspondence constraints that violate faithfulness. "Faithfulness" refers to the "elsewhere" case, in which lexical A corresponds to observed A.

8 Note that states 5, 9, and 13 are "pinch points" in the automaton: all paths through the automaton from states before 5 to states after 5 must pass through state 5, and similarly for states 9 and 13. In graph theory such nodes are called *articulation points*, and the fragments of graphs between them are termed *blocks*; there is a simple algorithm for splitting a graph into blocks (see Gibbons 1985: 25) that could be used to decompose a finite-state automaton into parts corresponding to some phonological constituents.

9 An example of this style of computation from antiquity is the *sieve of Eratosthenes*, a method for finding all the prime numbers up to n by striking out all the multiples of i from 2 to vn. The candidate set is all the integers from 2 to n, and the constraints are (expressed informally) "is not a multiple of 2," "is not a multiple of 3," etc.

10 Although this is an unpublished MA thesis, it is precisely what is needed here: a detailed Optimality Theory analysis of a nontrivial range of phonological phenomenon in Spanish, including syllabification and phonotactics.

11 Even though syllables must not have codas!

12 Since ONSET and SONSEQ are unranked, it is impossible to say which of them is fatal to forms which violate both of them.
13 The start and end of the string cannot be simply equated with the first and last input segment, because we allow for the possibility of leading and/or trailing empty constituents.
14 Connectionism is not the first framework to try to replace symbols and rules by circuits or networks of elementary processing elements. For a period in the 1960s, Stratificational Grammar attempted to encode grammar as switching circuitry: for a phonological overview, see Sommerstein (1977: Chapter 4). Systemic Phonology (e.g. Halliday 1992) has some similarities, though in this case the metaphor of a railway freight yard (Hockett 1958) is more apposite.
15 Connectionist research was also greatly advanced at that time by two important technical innovations: (i) adding a nonlinear compression function to the output of each unit, which enabled multi-layer networks to be developed that could compute various complex, non-linear mappings, and (ii) the invention of the back-propagation algorithm for adjusting unit weights, enabling networks to be trained by exposure to data.
16 Larson's nine-point hierarchy does not include mid vowels, so for completeness here we assign glides and high vowels an inherent sonority of 7, mid vowels 8, and low vowels 9.
17 This remark should be qualified by the observation that speech recognition system performance depends on many factors, including the vocabulary size, number of speakers, nature of the required output and so on. Nevertheless, it is noteworthy that in the competitions between state-of-the-art research systems organized by the US National Institute of Standards, HMMs and other non-connectionist statistical methods predominate, as they do in commercial systems for dictation transcription by desktop PCs.
18 http://ispell-fi.sourceforge.net/
19 This is a *good* thing. All theories are wrong, to some extent: to make progress, it is helpful to know *how* wrong, and *where*.

19 Using Psychological Realism to Advance Phonological Theory

MATTHEW GOLDRICK

1 Overview

Following its introduction by Sapir (1933), the term "psychological reality" has provoked intense reactions from within linguistics as well its neighboring disciplines. Discussions have been particularly heated since the rise of generative grammar, whose proponents made quite strong claims regarding the relationship of theoretical concepts from linguistics to the internal cognitive mechanisms underlying the acquisition and processing of sound patterns. For example, *The Sound Pattern of English* is asserted to be "a hypothesis concerning the actual internalized grammar of the speaker-hearer" where grammar refers to "a system which is used in the production and interpretation of utterances (Chomsky and Halle 1968: 4)." Although this perspective is by no means universally adopted by phonologists, its dominance in linguistics since the mid-twentieth century reflects a major conceptual shift from previous perspectives on the study of language. As noted by Anderson (1985: 6; emphasis in original):

> Traditionally, linguists have assumed that their concern was the study of *languages*, taken as (potentially unlimited) sets of possible sentences (or utterances, etc.) forming unitary and coherent systems. Gradually, however, the emphasis in research has shifted . . . to the properties of *grammars*, in the sense of systems . . . which specify the properties of the (well-formed) sentences in such a system.

This shift, coupled with the claim that linguistic systems define a capacity (or competence) possessed by *individual* speakers of a language, has focused attention on the correspondence between constructs from linguistic theories and the

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

cognitive systems of individual speaker-hearers. *Psychological realism* views such correspondences as cornerstones of linguistic research – both in terms of empirical practice and theory development. This chapter considers the content and import of this approach in competent, adult individuals (for discussion of language acquisition, see Demuth, this volume). Three core issues are considered:

- *What is psychological realism?* Psychological realism adopts a cognitive psychological perspective to explain human linguistic behavior. This offers a functional-level account of how different components of the human cognitive system interact to yield particular behaviors.
- *Why is psychological realism critical for linguistic research?* Human behavior always reflects the interaction of multiple cognitive components. Without making explicit (and empirically justified) assumptions about the nature of these interactions, we cannot correctly draw inferences about the structure of the cognitive system. The perils of failing to specify these assumptions will be illustrated using well-formedness judgments.
- *How can psychological realism help resolve theoretical issues in linguistics?* If we take seriously the need to articulate the functional architecture underlying specific tasks, we can better understand the import of behavioral data. This can help resolve outstanding theoretical questions such as the nature of the relationship between lexical and grammatical knowledge.

2 What is Psychological Realism?

2.1 The Structure of Psychological Theories

Psychological realism adopts the theoretical perspective of cognitive psychology to understand language-related behavior. (Note that this is by no means the only perspective on human psychological capacities; see, e.g. van Gelder 1998.) It explains language behavior as the coordinated interaction, in real time, of a set of more primitive capacities or *functions* that map between inputs and outputs.

2.1.1 Functional Explanation Cognitive psychological theories aim to specify the psychological capacities of individuals. Capacities are the regularities that govern the behavior of the cognitive system. These regularities are *lawlike* in that given certain precipitating conditions (environmental or internal), the system will exhibit certain (behavioral or internal) manifestations (Cummins 1983). The term "function" is used to evoke the idea of a precisely specified relation or mapping between starting configurations of the system (i.e. precipitating conditions) and ending configurations (manifestations). When describing the cognitive system at this level of description, there is no specification of *how* this mapping is accomplished. Our theory simply specifies that given certain inputs, a certain distribution of outputs will be produced. (Note this mapping could be deterministic or stochastic.) This corresponds to Marr's (1982) *computational* level of analysis of

cognitive systems. The discussion here follows Smolensky (2006) in referring to this as the *functional* level of description.

To illustrate this level of analysis, many psycholinguistic theories of speech perception assume two broad functional stages are involved in the perception of single spoken words (McClelland, Mirman, and Holt 2006). The first stage, *prelexical processing*, takes as input a relatively fine-grained representation of acoustic information (e.g. acoustic features) and produces as output a prelexical representation elaborating the linguistic structure of the acoustic input (e.g. by specifying segmental and prosodic structure). The second stage, *lexical processing*, uses this prelexical representation to retrieve a lexical representation of the utterance (e.g. a unitary whole-word representation; this is used to access semantic and syntactic information).

Although this level of description does not specify *how* a mapping is computed, claims stated at this level are contentful statements about the psychological organization of speakers. First, as anyone who has attempted to construct a generative grammatical analysis can tell you, it is no trivial matter to precisely specify a function that maps a large set of inputs to the correct set of outputs. Second, functional analyses represent a critical initial step in the pursuit of reductionist accounts of behavior. Functional theories take a complex capacity such as "the ability to perceive the meaning of single spoken words" and decompose it into simpler capacities – for example, "mapping acoustic signals onto phonemes," "retrieving words matching the perceived phonemes," and "retrieving the meaning of words." (The hope is that this reductionist procedure will terminate in simple capacities which can be realized as neural computations; see below). That such claims are taken to be contentful is clear from many psycholinguistic studies aiming to distinguish theories with contrasting decompositions of complex capacities. For example, with respect to the capacity of perceiving single spoken words, other psycholinguistic theories of perception have proposed that there is no explicit prelexical stage intervening between acoustic signal processing and meaning retrieval (e.g. Gaskell and Marslen-Wilson 1997).

2.2 *The Role of Capacities in Real-time Linguistic Behavior*

Cognitive psychological theories aim to account for linguistic behavior. An essential component of such theories is therefore specifying how functions are utilized in the performance of specific behaviors or tasks. For example, what functions are utilized in an auditory lexical decision task (where a participant must decide if a string of sounds is a lexical item or not)? Within the general framework outlined above, it is generally assumed that performance in this task is related in part to the outcome of lexical processes; participant responses reflect at least in part whether a word representation is or is not successfully retrieved for the input. Within the architecture above, this means that prelexical processes are also engaged; sensory input cannot influence lexical processing without the mediation of prelexical processes. Subsequent to lexical processing, there must also be decision processes that allow the hearer to make a word/non-word

response (see Ratcliff, Gomez, and McKoon 2004, for a recent review of such models in the context of lexical decision tasks using printed words).

It is critical to note that within psychological theories functions, like human behavior, exist in real time; they have temporal extent. For example, Palmer and Kimchi (1986: 40) define psychological capacities as "informational events" consisting of "the *input information* (what it starts with), the *operation* performed on the input (what gets done to the input) and the *output information* (what it ends up with)" (emphasis original). Rather than speaking broadly of a function as an a temporal specification of a relation or function between input and output, this approach makes the stronger assumption that the capacity literally starts with the input at some time and after some distinct period actively *produces* the output. Likewise, the interaction of these capacities is inherently temporal. According to Palmer and Kimchi, decomposition of a complex capacity is specification of a set of informational events plus "the temporal ordering relations among them that specify how the information 'flows' through the system of components" (p. 47). That is, the primary task of structuring the interaction of simpler capacities is specifying their temporal relationships.

2.3 *Psychological vs. Algorithmic or Neural Accounts*

Psychological theories, stated at the functional level of description, do not offer complete accounts of the cognitive system. Ultimately a complete theory must specify not only what functions are but (i) how they are computed and (ii) how they are realized physically by the nervous system (Marr 1982). Accounts that address these two issues are sketched below.

2.3.1 The Algorithmic Level The algorithmic level of description is an abstract, computational characterization of the process which satisfies the function specified at the higher level of description. That is, if the cognitive system is placed in the appropriate initial configuration, the algorithm will place the system in the desired ending configuration.

For example, the lexical function described above has been instantiated within spreading activation networks (e.g. McClelland and Elman 1986; Norris, McQueen, and Cutler; 2000). In these networks, one set of processing units instantiates prelexical speech sound representations (e.g. there is a unit corresponding to initial /d/, another for initial /t/, etc.). Another set of units instantiates the lexical representations (e.g. there is a unit for DIG, another for DOT, etc.). The function specified above is realized by connection weights that allow activation to flow between these two levels of representation. When the word-initial input /d/ is provided to the network by imposing a certain pattern of activation on the prelexical units (e.g. activating /d/ but not /t/), activation will automatically flow along these connections to the appropriate output units (e.g. activating DIP but not TIP). This process is entirely mechanical; given an input, the network will automatically produce (via spreading activation) the output that satisfies the function specified above.

2.3.2 Neural Level Accounts Of course, a complete account of the cognitive system cannot stop at the algorithmic level. Human cognitive systems are ultimately realized by neurobiological structures and processes. Therefore, at the lowest level of description – the neural level – the algorithms specified at the preceding level are implemented in terms of neural systems. For example, the process of accessing lexical representations from acoustic input has been argued to be instantiated by brain structures in the vicinity of the temporal-parietal-occipital junction (see Hickok and Poeppel 2000 for a recent review). A neural specification of the spreading activation networks would have to detail how the algorithm specified above (i.e. abstract prelexical as well as word-sized processing units; activation flow among these units) is instantiated in these neuronal assemblies.

2.3.3 Psychological Accounts Are Not Algorithmic or Neural Accounts Cognitive psychological explanations are, in general, stated in terms of capacities; that is, they are *functional* level explanations. They do not typically address the algorithmic (much less the neural) realization of capacities characterized at the functional level. For example, following the general reductionist strategy of functional accounts, cognitive psychological theories account for complex behavior in terms of the interaction of (relatively) simple capacities (Cummins 1983; Palmer and Kimchi 1986). These simple component capacities are assumed to be physically embodied (i.e. algorithmically and neurally realized) but the details of how this occurs are typically not spelled out. "In reality, most IP [information-processing or cognitive psychological] theorists give, at best, a rather vague, verbal description of the input-output characteristics of the components . . . unfortunately, simulations [algorithmic implementations] are seldom actually done . . ." (Palmer and Kimchi 1986: 53–54).

That is not to say that psychological (or linguistic) theorizing categorically avoids other levels of description. In particular, many theories are at least partially specified at both the functional and algorithmic levels. For example, as discussed above, the processing of monosyllabic monomorphemic forms has been computationally implemented by a spreading activation network (note, however, the limited range of inputs this specific algorithm can process). However, such work is the exception rather than the rule. In most psycholinguistic theories, many processing components are wholly unanalyzed algorithmically (as noted by Palmer and Kimchi above). The situation is far worse with respect to the neural level of description. As far as I am aware, no theories of language-related capacities have attempted to specify the physically instantiated neurobiological processes that realize cognitive functions (although many have investigated the neurobiological structures associated with linguistic capacities). For example, although connectionist research has attempted to specify algorithms that are broadly compatible with neuronal computational principles, it is still extremely abstract relative to actual neurobiological mechanisms (Smolensky 2006).

Psychological realism is therefore like most research in linguistics; it adopts a functional level approach to understanding human behavior. A critical issue for

any functional theory is *realizability*: how is the functional level description instantiated algorithmically and, ultimately, neurally? These issues are critical, as they address the *physical* reality of theoretical constructs. If no algorithm can be specified that instantiates a hypothesized function, or if there is no way to realize that algorithm neurobiologically, the functional level description becomes significantly less plausible. But (contra authors such as Linell 1979) these issues are distinct from specifying a psychological (functional level) account of linguistic behavior.

2.4 *Linking Linguistic and Psychological Theories*

As noted above, linguistic theories are also, in general, functional level theories. They decompose complex linguistic knowledge into a set of simpler functions (e.g. syntactic vs. phonological components of the grammar). Although research in computational phonology aims to specify algorithms that compute grammatical functions (see Coleman, this volume, for further discussion), the typical linguistic analysis does not consider how it is computationally (much less neurally) implemented.

Psychological theory can enrich such theories by providing a framework for thinking about how these linguistic functions are deployed during behavioral tasks. Like linguistic theories, cognitive psychological theories explain language behavior as the coordinated interaction of a set of more primitive functions that map between inputs and outputs. Unlike many linguistic theories, psychological accounts are situated within specific behavioral tasks and in real time. This enables psychological theories to make predictions (that can be confirmed or refuted) for behavioral experiments. Linking a linguistic theory with a psychological theory allows the linguistic theorist to draw upon this rich body of evidence to inform their theory.

Making such connections is facilitated by the use of functional level descriptions in each tradition. However, difficulties can arise due to contrasting assumptions regarding the specificity of linguistic knowledge. Most linguistic theories aim to characterize capacities common to all linguistic behaviors; in contrast, many psychological theories aim to characterize the capacities involved in particular sets of behaviors.[1] For example, the psychological theory discussed above concerns the relationship between various levels of sound structure in speech perception; it makes no claims regarding speech production. In contrast, a typical linguistic theory would attempt to characterize the general relationship between levels of sound structure representation – a relationship that subserves perception, production, acquisition, well-formedness judgments, and so on. For example, Jakobson (1941: 92) claims "the *same laws* of solidarity (emphasis mine)" underlie child language, aphasia, and typological sound structure patterns. Chomsky and Halle (1968) assume the (singular) grammar is "*a* system used in the production *and* interpretation of utterances" (p. 4; emphasis mine).

A critical issue in psychological realism is therefore establishing how components of linguistic theories link up to components of psychological theories. Without such links, it is impossible for linguistic theories to use psychological theories

to help draw inferences from behavior. The nature of such links cannot be established a priori. Some theories have assumed relatively direct connections between components of linguistic theories and psychological mechanisms (e.g. Goldrick and Daland 2009). However, linguistic theories making differing assumptions regarding the specificity of knowledge may necessitate more complex relationships with psychological theories. For example, the phonological component of the grammar may be distributed over many distinct psychological capacities (specific to memory, language production, perception, etc.).

In spite of such complexities, the establishment of these links is imperative for linguistic theorists that wish to make use of behavioral data. If a linguistic theory is not situated within specific tasks that occur in real time it cannot be informed by data from online behavioral tasks. Psychological theories provide an appropriate set of linking hypotheses licensing such inferences.

3 Why is Psychological Realism Critical for Linguistic Research?

As discussed above, psychological theories account for behavior in any given task through the complex interaction of *many* simple capacities. This point has been noted by many authors; for example, Chomsky (1980: 188) writes "the system of language is only one of a number of cognitive systems that interact in the most intimate way in the actual use of language." Such interactions clearly complicate the interpretation of behavioral data. When assessing behavior, it is not sufficient that we richly articulate a theory of the cognitive component of interest (e.g. the "phonological grammar"). We must also specify how this component interacts with other relevant cognitive process to produce the behavior(s) of interest. As discussed by Caramazza (1986: 47):

> observations do not carry on their sleeves signs indicating whether or not they constitute relevant evidence in some domain of investigation. An especially important point is that a specific set of observations . . . will assume evidential status with respect to some model only if we are able to provide adequate arguments . . . to explicitly link the type of observations in question to the component or components of processing being investigated.

Using behavioral data to inform theories therefore *requires* "a sufficiently detailed model of the cognitive systems of interest to guide the search for richly articulated patterns of performance" (Caramazza 1986: 66). Until this has been specified to at least some level of detail, we cannot establish (to use Caramazza's term) the "evidential status" of behavioral data. These concerns are by no means unique to cognitive psychologists. For example, Fodor (1981: 200) notes that "[a]ny science is under the obligation to explain why *what it takes to be* data relevant to the confirmation of its theories *are* data relevant to the confirmation of its theories" (emphasis original).

In this section, we examine a critical domain of behavioral data – word-likeness judgments – where the inference from data to the structure of the cognitive system has been impaired by the failure to consider how multiple cognitive components interact to produce behavior. The particular domain we focus on here is phonotactic knowledge – our knowledge that certain combinations of phonological structures are dispreferred relative to others. First, the role of such knowledge in generative theories is briefly reviewed. We then critically review the data from word-likeness judgments.

3.1 Generative Models of Phonotactic Knowledge

3.1.1 Phonotactic Knowledge Evidence from a wide array of linguistic behaviors suggests that our cognitive systems are structured in such a way as to disfavor certain combinations of phonological structures relative to others (note these preferences are to a certain degree language specific, but general patterns are found across languages). For example, although English words contain both /p/ and /l/, a phonological string with an initial cluster /lp/ will be disfavored over a string with an initial /pl/. The dispreference for certain structures may manifest itself in a number of different ways, including: native speaker judgments of acceptability (a string with initial /lp/ is judged to be a poor English word); statistical under-representation (or absence) in corpora (there are no words in English with initial /lp/); difficulties in memory, perception, and production (for English speakers, it is difficult to recall, perceive or produce initial /lp/ clusters). *Phonotactic knowledge* concerns how cognitive functions are structured so as to yield these behaviors. The discussion here assumes that phonotactic knowledge distinguishes among forms in terms of their degree of well-formedness. Favored structures are well-formed; disfavored structures are ill-formed.

Note that phonotactic knowledge may allow speakers to disfavor structures to varying degrees. For example, consider the fricatives /f, v, h/ in word-final position. In English, sequences with /h/ in this position are completely absent, while /f/- and /v/-final sequences (laugh, live) are attested – suggesting that word-final /h/ may be strongly disfavored by English speakers' phonotactic knowledge. However, although both are attested, /v/ is much less frequent than /f/ in this position; /v/ is found in fewer words and occurs with a lower frequency in running speech. This suggests that relative to /f/, word-final /v/ may be disfavored to some degree. Phonotactic knowledge may therefore also make gradient distinctions in well-formedness.

3.1.2 Generative Phonological Models There is a long history in linguistic theory of seeking to develop theories of phonotactic knowledge. "Whereas traditional phonology generally gives rules for articulating all sounds . . . and stops there," Saussure (1916: 51) writes, "combinatory phonology limits the possibilities and defines the constant relations of interdependent phonemes." The dominant theoretical framework since the time of Chomsky and Halle (1968) has been

generative grammars. A generative grammar specifies a relation mapping a set of underlying phonological structures to a set of surface phonological structures (Smolensky, Legendre, and Tesar 2006; note the discussion here assumes a probabilistic formulation of this relation). Such grammars model phonotactic knowledge by specifying probability distributions over the set of surface phonological structures; well-formed structures are assigned higher probability than ill-formed structures.

Most work in this tradition has focused on a binary distinction between legal vs. illegal strings, characterizing the latter as categorically ill-formed and the former as categorically well-formed (see Goldsmith 1995, for a review).[2] Categorically well-formed structures are all generated by the grammar with equal probability (i.e. they are all equally well-formed); categorically ill-formed structures are never generated by the grammar (i.e. they are all equally ill-formed). For example, if we believe that the phonotactic knowledge of an English speaker specifies that /h/ is ill-formed in word-final position, we can represent this within a generative grammar in two ways. First, we can alter the probability distribution over underlying structures. For example, we could ban underlying representations containing /h/ in this position (e.g. morpheme structure constraints). A second mechanism involves altering the structure of the grammatical function. For example, we can ban any mapping (as defined by a set of rules or a constraint ranking) that allows /h/ to be generated in this position.

More recently, interest has grown in modeling gradient distinctions in phonotactic knowledge. This has been most frequently modeled in terms of generation probability. Less well-formed structures have lower generation probability than more well-formed structures. For example, if we believe that the phonotactic knowledge of an English speaker specifies that /v/ is disfavored in word-final position relative to /f/, we can represent this within a generative grammar by (a) assigning underlying representations containing /v/ in this position lower probability than comparable representations containing /f/ (e.g. Frisch, Pierrehumbert, and Broe 2004, assume a similar gradient constraint on roots in Arabic) and/or (b) assigning a lower probability to any mapping resulting in /v/ rather than /f/ in this position. A variety of formal mechanisms have been specified to assign probabilities to mappings: within derivational theories by associating probabilities to rule applications (e.g. Coleman and Pierrehumbert 1997); or within Optimality Theoretic (Prince and Smolensky 1993) approaches by assigning probabilities to constraint rankings (e.g. Boersma and Hayes 2001) or output candidates (e.g. Coetzee 2006; Hayes and Wilson 2008). Further extensions to these formal mechanisms can allow generative grammars to represent gradient distinctions in well-formedness among unattested structures (e.g. for an English speaker, assigning varying probabilities to the generation of unattested initial clusters /fn/ and /zg/; Davidson 2006b).

The discussion below examines the influence all of these various types of well-formedness distinctions on behavior: categorical distinction between attested and unattested strings (e.g. in English, final /f/ is possible, but final /h/ is not); gradient distinctions among attested strings (e.g. final /v/ is disfavored relative

to final /f/); and gradient distinctions among unattested strings (e.g. for an English speakers, varying preferences for initial /fn/ vs. /zg/). In each case, these well-formedness distinctions are part of the (hypothesized) mental knowledge that underlies behavior. Note that certain theories use objective measures (e.g. relative frequency or probability of structures in corpora) to estimate these mentally represented distinctions in well-formedness. It is important not to confuse these two notions; although the objective measures are used to *estimate* phonotactic knowledge, it is the mentally represented distinctions in well-formedness that are casually involved in producing behavior.

3.2 *Inferring Phonotactic Knowledge from Well-formedness Judgments*

To develop models within the various formal frameworks discussed above, researchers have drawn inferences concerning our knowledge of well-formedness distinctions based on various types of empirical data. Much of this work relies on categorical judgments of acceptability or possibility of various forms (e.g. "Is /zɑh/ a possible English word?"). (These judgments may be systematically organized and codified in a written grammar.) The patterns identified within this set of judgments (e.g. the tendency to judge forms with word-final /h/ as unacceptable) then inform the construction of generative grammatical models (e.g. the postulation of morpheme structure constraints banning /h/ in this position). Judgments can also inform models by providing a test of their predictions. For example, suppose a model classifies a form as categorically well-formed. If the form is judged to be unacceptable, this provides some evidence against the model; if it is acceptable, the evidence is consistent with the model.

Below, we briefly review three issues with using this type of data to inform grammatical theories. The first is purely methodological: more quantitative methods are required to accurately assess judgments. We then turn to two more substantive issues with this work. These issues reflect the failure of this type of research to consider psychological realism – how the cognitive system is structured so as to yield judgments. As discussed below, these issues are likely to lead to errors in inferring well-formedness distinctions from behavior.

3.2.1 Issue 1: Quantitative Analysis of Behavior Although the collection of judgments and identification of patterns within them is done with great care and precision, it often does not do complete justice to the complexity of the underlying behavioral data from judgments. Frequently (as in many written grammars) acceptability judgments are codified as binary distinctions in well-formedness. (This reflects the assumption that acceptability judgments are a more or less direct reflection of phonotactic knowledge – a conflation of behavior with the mental representation of degrees of well-formedness.[3]) Critically, this binary categorization obscures a fair amount of variation and gradience in participant responses (see also Bard *et al.* 1996 for discussion).

The problem of limiting judgments to binary distinctions is addressed by studies that utilize more quantitative assessments of participants' judgments. For example, Greenberg and Jenkins (1964) asked participants to provide a numerical estimate of how far a stimulus word was from English (a technique called free magnitude estimation). Other studies have used rating scales: for example, Vitevitch *et al.* (1997) asked participants to rate stimuli on a scale from 1 (bad example of an English word) to 10 (good example of an English word). The above scales ask participants to judge how word-like a given stimulus is; other studies focus more on "acceptability." For example, Berent and Shimron (2003) asked participants to rate, from best to worst on a five-point scale, how a stimulus sounds. Bailey and Hahn (2001) asked participants to rate how typical a stimulus sounds (on a scale from 1 to 9).

Studies using these types of tasks report that participants reliably judge structures that are classified by grammatical models[4] as ill-formed as being less acceptable or word-like than structures classified as well-formed (e.g. Arabic: Frisch and Zawaydeh 2001; Cantonese: Kirby and Yu 2007; English: Greenberg and Jenkins 1964; Pierrehumbert 1994; Hebrew: Berent and Shimron 1997; Hindi: Ohala 1983; Tagalog: Zuraw 2007; Turkish: Zimmer 1969).

Cross-linguistic research has also documented that participants' ratings of attested items correlate with relative degrees of well-formedness (as predicted by grammatical models incorporating gradient distinctions). Work using judgments to test gradient models of phonotactic well-formedness has a long history; it was a key motivation behind Greenberg and Jenkins' (1964) seminal word-likeness study. They in fact found that English speakers' judgments were graded, consistent with the mental representation of gradient distinctions in well-formedness. Subsequent studies utilizing multiple methodologies in many languages have also documented gradient distributions in participants' judgments. In studies examining categorical judgments by English speakers (e.g. "is this a possible English word?"), mean ratings across participants or items are gradient (e.g. Coetzee 2008; Coleman and Pierrehumbert 1997; Dankovicova *et al.* 1998; Scholes 1966). Gradient distinctions are also found when individual English speakers make use of scales on each item (e.g. Bailey and Hahn 2001; Dankovicova *et al.* 1998; Frisch *et al.* 2000; Ohala and Ohala 1986; Shademan 2006, 2007; Vitevitch *et al.* 1997). Similar results have been reported in a diverse set of other languages (Arabic: Frisch and Zawaydeh 2001; French: Perruchet and Peerman 2004; Korean: Lee 2006; Tagalog: Zuraw 2007).

Although quantified observations allow for a more nuanced picture of the distributional properties of behavioral responses, the precise connection of these observations to the underlying phonotactic knowledge is not entirely clear. Scant attention has been paid to how phonotactic knowledge is deployed in real time to yield judgments of well-formedness (see Schütze 1996 for detailed discussion of similar issues in grammaticality judgments). Without clarifying, to some degree of approximation, how this task is performed, it is difficult to determine the precise implications of these results. The sections below consider two areas where this lack of clarity could lead to incorrect inferences regarding phonotactic knowledge.

3.2.2 Issue 2: Dynamic Weighting of Multiple Factors in Judgments It is likely that, as in other decision tasks, word-likeness or acceptability judgments are not a pure reflection of well-formedness; judgments most likely reflect a combination of factors. Following this latter assumption, most research assumes judgments reflect both phonotactic well-formedness as well as similarity to existing lexical items (e.g. Bailey and Hahn 2001; Shademan 2007).[5] These two factors are conceptually distinct. There are forms that are unattested but well-formed (e.g. *hing*). Measures of similarity to existing lexical items will be sensitive to such absences, whereas a pure measure of phonotactic well-formedness would not. Empirical results suggest that these two factors independently contribute to judgments. When similarity to existing items is controlled, forms that are classified as ill-formed are still judged less word-like than those classified as well-formed (e.g. Arabic: Frisch and Zawaydeh 2001). For English nonwords classified as relatively ill-formed, word-likeness judgments correlate with degrees of well-formedness, not similarity to existing lexical items (e.g. Frisch, Large, and Pisoni 2000; Coetzee 2008). Finally, regression analyses can examine the degree to which judgments are influenced solely by degrees of well-formedness, similarity to existing items, or some combination of the two factors. Such analyses show well-formedness exerts an independent effect on judgments (e.g. Cantonese: Kirby and Yu 2007; English: Albright 2009; Bailey and Hahn 2001; Shademan 2006, 2007). (Note that these two factors do interact in judgments; see Shademan 2007 for discussion.)

Although these studies have taken into account the influence of multiple factors on participant judgments, little work has addressed the possibility that these influences are not static. A basic finding in psychology and psychophysics across many paradigms and domains is that decision processes are dynamic – i.e. sensitive to the context in which they are presented (see Vickers and Lee 1998, for a general review). It is therefore unlikely that a simple static function maps judgments to internally represented well-formedness distinctions. There are a number of examples of such contextual effects in psycholinguistics. Consider the well-studied task of lexical decision (where participants judge whether an auditory or visual stimulus corresponds to a word in one of their languages). Judgments are influenced by the composition of non-word fillers (e.g. more word-like fillers tend to slow responses), the proportion of high vs. low frequency targets, and repetition of target items (see Ratcliff, Gomez, and McCoon 2004 for a recent review). In same-different judgments with non-word auditory stimuli (e.g. is "bep" different from "mep"?), Vitevitch (2003) found the degree to which filler items are composed of real lexical items leads to shifts in the relative weighting of different factors. The use of mostly word fillers leads to a greater influence of similarity to existing lexical items, whereas mostly non-word fillers leads to a greater weighting of well-formedness. Critically, recent results suggest that word-likeness judgments are similarly sensitive to properties of the experimental context in which the judgment is given. Shademan (2006, 2007) reported changes to the relative weighting of similarity to lexical items and well-formedness in judgments depending on whether the stimulus set contains both words and nonwords or is composed of nonwords only. Although her results are not clearly consistent across

analysis methods, some results suggest that when words are excluded from the stimulus set, participants' judgments tend to more strongly reflect well-formedness. Finally, Shademan (2007) reported that relative to those of young adults, the judgments of healthy older individuals are more sensitive to similarity to existing lexical items. These findings provide some preliminary support for the claim that word-likeness judgments are context dependent – just like decisions in many other cognitive domains.

If factors other than well-formedness exert a *variable* influence on word-likeness judgments, utilizing such judgments to inform linguistic theories becomes less straightforward. For example, suppose including real words as well as nonwords within in an experiment causes individuals' judgments to reflect more strongly similarity to existing lexical items. This may cause judgments to be less sensitive to gradient distinctions in well-formedness. Without awareness of such effects, we might mistakenly conclude that well-formedness distinctions are mentally represented in a more categorical fashion than they actually are. It is imperative for researchers using word-likeness judgments to investigate how contextual factors influence judgment performance. This will allow for development of more complete models of the judgment process involved in this task (in much the same way that extensive investigation of context has enriched theories of lexical decision performance).

3.2.3 Issue 3: The Interface of Judgment Processes with Other Cognitive Processes As repeatedly emphasized above, any behavior reflects the complex interaction of multiple psychological capacities. Word-likeness judgments are no different; in order to make a word-likeness judgment, one must perceive the acoustic structure of the form, assign a phonological parse to it, and so on. However, such relations have been left largely unspecified in the literature. The failure to articulate how judgments are situated within the cognitive system could lead to misinterpretations of behavioral data.

Suppose that in perception the phonological grammar corresponds to some specific cognitive process. Judgment processes respond to the output of this grammatical component of the cognitive system. However, this does not mean that judgments are influenced solely by the grammar; this component is but the last link in a causal chain of processes mediating the stimulus and the word-likeness judgment.

Suppose we present an English speaker with two stimuli: one containing a word-final /h/, the other a word-final /s/. We then ask them to characterize the relative acceptability of the two stimuli. One possible result is that they will judge the form with /h/ as being less acceptable. Does this imply that their grammatical processes encode word-final /h/ as less well-formed than word-final /s/? Not necessarily. It is possible that some well-formedness effects emerge in earlier stages of processing. For example, ill-formed sequences may be corrupted or distorted by relatively early perceptual processes (e.g. when exposed to an /h/ in word-final position, an English listener may have difficulty perceiving the intended sound). There is some empirical support for such a possibility; studies

in several languages have shown that in fairly basic perceptual tasks hearers have difficulty perceiving categorically ill-formed stimuli – such that their perception is distorted towards a well-formed percept (e.g. Dupoux *et al.* 1999; see below for further discussion). This "repair" of the perceptual input may result in a corrupted or distorted representation (i.e. it may be an incomplete repair – a poor exemplar of the well-formed percept). This distorted input to grammatical processes might then cause hearers to judge a stimulus as less word-like. The distorted quality of the input – rather than the mental representation of well-formedness between /h/ and /s/ – would give rise to the judgment of word-final /h/ as less acceptable than word-final /s/.

Alternatively, suppose the speaker judges word-final /s/ and /h/ to be equally acceptable. Does this imply that their grammatical processes encode word-final /h/ and /s/ as being equally well-formed? Not necessarily. Suppose that perceptual repairs do not result in distorted inputs but instead produce completely well-formed representations. Under this scenario, the repair converts a categorically ill-formed representation to a categorically well-formed representation (see below for further discussion). A word-final /h/ stimulus is therefore transformed by perceptual processes into a completely well-formed word-final /s/. Even if the participant's grammar encodes word-final /h/ as less well-formed, this perceptual transformation prevents grammatical processes from influencing judgment behavior.

We cannot draw a simple, direct connection from judgments to the mental representation of degrees of well-formedness. One means of attempting to circumvent this issue is to assume that the phonological component of the grammar is in fact distributed across multiple psychological capacities. Thus, the influence of well-formedness on more basic perceptual processes still has implications for the structure of the grammar. This may be a promising approach, but it would require a great deal of specification to become a plausible hypothesis. For example, returning to the issue raised above, if the multiple psychological processes encoding the grammar are each subject to independent contextual variation, the problem of relating judgments back to grammatical structure becomes many orders more complicated. Assuming the grammar is distributed over multiple psychological processes does not eliminate the need to consider the functional architecture of these processes; if anything, it makes such issues even more critical.

3.3 The Perils of Avoiding Psychological Realism

The common use of binary classifications to characterize judgments is a clear methodological issue in many studies. But the exclusive use of well-formedness judgments, disconnected from any theory of human language processing, has much deeper flaws that are likely to yield incorrect inferences regarding the nature of our knowledge of well-formedness.

In order to correctly draw inferences from online behavioral data, linguistic theories must be situated within specific tasks that occur in real time. Current research into well-formedness judgments has not adequately addressed this problem; neither the judgment process itself nor its interface with other cognitive

systems have been well articulated or empirically explored. As discussed above, this lack of clarity is likely to lead to not only imprecise but also faulty conclusions regarding the nature of phonotactic knowledge. Because all linguistic data (including distributional data in corpora or dictionaries) ultimately arises from online behavior, interpretation of such data requires serious consideration of the symphony of coordinated cognitive functions that gives rise to that behavior. Of course, it is highly likely that current theories will be incomplete in many respects; their cognitive functions will be less than completely specified, their interactions only partially spelled-out. But even a partial account will provide purchase on many of the issues identified above. It will provide *some* justification for linking behavioral observations to the cognitive component(s) of interest.

This is not to say that well-formedness judgments bear *no* relation to phonotactic knowledge. It is highly likely that phonotactic knowledge does contribute in some way to well-formedness judgments. The critical issue is that without specifying the nature of this contribution in greater detail, we cannot draw inferences from behavioral data. We have no basis for arguing that variation in judgments reflects the function(s) of interest rather than some other function involved in the judgment task.

4 How Can Psychological Realism Help Resolve Theoretical Issues in Linguistics?

4.1 Utilizing Psychological Realism

To establish links between behavioral data and the structure of particular cognitive functions, psychological research adopts two basic strategies. One is through detailed examination of particular tasks. Researchers adopt theories of a particular function and its interaction with other functions; empirical research is then used to test and refine these theories (both in terms of the structure of functions as well as their interactions). A second approach is to seek converging evidence from a wide variety of tasks. Many cognitive functions are typically assumed to be utilized in a wide variety of tasks. For example, psycholinguistic theories of speech perception typically assume that there is a common set of cognitive functions engaged in all perceptual tasks. At some level this is obviously true; for any perceptual task, the hearer must at a minimum perform some analysis of the speech sounds (otherwise there would be no stimulus to perform the task on!). Researchers can rely on these functional commonalities to draw on performance across a wide range of tasks to constrain theories of specific cognitive functions.

The sections below examine how these strategies have been used to examine the nature of phonotactic knowledge in both speech perception and production. The first step is to establish that phonotactic knowledge is in fact utilized in both perception and production. To support this claim, the functional architecture of core processes involved in perception and production is outlined. These architectures characterize how phonotactic knowledge is deployed online in these behavioral

tasks. Converging evidence from a variety of behavioral tasks that engage these core processes is then reviewed. This body of work provides support for the hypothesized functional architecture's claim that phonotactic knowledge plays a critical role in these tasks.

Armed with some validation of our linking assumptions, we can then turn to more detailed questions regarding the nature of phonotactic knowledge. In the course of reviewing data from perception and production, we touch on two types of results that bear on two key issues in linguistic theory: the relationship between grammar and lexicon and the gradient vs. categorical nature of phonotactic knowledge.

Within the generative tradition systematic aspects of phonological knowledge (e.g. phonotactics) have typically been reflected by the structure of the grammar. This is seen as a distinct component of the cognitive system from that of the (phonological) lexicon (Chomsky and Halle 1968). This specifies those phonological structures which correspond to lexical items of the language (i.e. words). However, this distinction has not been universally adopted, even with, the generative tradition. Recent work in linguistic theory has strongly emphasized the view that lexicon and grammar are highly intertwined, inseparable components of the cognitive system (e.g. Burzio 1996; Bybee 2001).

The second issue addressed below concerns the nature of phonotactic well-formedness. As discussed above, many generative grammatical models have focused on categorical distinctions in well-formedness. However, more recently there has been a growing interest in modeling gradient distinctions in grammatical knowledge. There has been considerable disagreement in the literature regarding whether the former or latter perspective best characterizes the nature of our linguistic knowledge (see, e.g. Newmeyer 2003, and associated commentaries in *Language* 81(1) for a recent discussion).

As discussed below, behavioral data – interpreted within functional frameworks for speech perception and production – can help resolve these questions. With regard to the first issue, cognitive functions encoding phonotactic knowledge are distinct from (but interact with) those encoding word-specific knowledge; this is consistent with the claim that lexicon and grammar are independent components of linguistic knowledge. Second, gradient distinctions in phonotactic well-formedness influence perception and production – suggesting phonotactic knowledge is not limited to categorical distinctions.

4.2 Phonotactic Knowledge in Speech Perception

4.2.1 Phonotactic Knowledge Within the Functional Architecture of Speech Perception As discussed above, many psycholinguistic theories of speech perception assume two broad functional stages in speech perception tasks involving stimuli up to the size of single spoken words (e.g. McClelland, Mirman, and Holt 2006). The first stage takes as input a relatively fine-grained representation of acoustic information (e.g. acoustic features) and produces as output a prelexical representation elaborating the linguistic structure of the acoustic input (e.g. by

specifying segmental and prosodic structure). The second stage uses this prelexical representation to retrieve a lexical representation of the utterance (e.g. a unitary whole-word representation <DOG> which is used to access semantic and syntactic information). In this discussion the first stage will be referred to as *prelexical processing* and the second as *lexical processing*. (Note: these two stages of processing may overlap and interact; see McClelland, Mirman, and Holt 2006; Norris, McQueen, and Cutler 2000, for discussion). This basic decomposition of the system into two distinct lexical and prelexical representations has received support from a variety of perceptual tasks (for recent reviews, see Gaskell *et al.* 2008; McQueen, Cutler, and Norris 2006).

In this two-stage framework, phonotactic knowledge is assumed to be encoded within prelexical processes; the speed and accuracy with which prelexical representations are activated reflects their phonotactic well-formedness. For example, with respect to categorical well-formedness, Dupoux *et al.* (1999) assume that categorically ill-formed sequences lack stored suprasegmental representations possessed by categorically well-formed sequences (specifically, demisyllables). The lack of such stored representations leads to greater errors and slowed processing times on ill-formed sequences. With respect to gradient distinctions in phonotactic well-formedness, Luce *et al.* (2000; see also Vitevitch and Luce 1999) assume that representational units within relatively well-formed sequences facilitate one another – allowing for more rapid and accurate retrieval for more vs. less well-formed structures (for discussion of similar mechanisms, see Newman, Sawusch, and Luce 1999; Norris *et al.* 2000).

This functional architecture predicts a number of behavioral manifestations of phonotactic well-formedness. This section focuses on sub-word perceptual tasks in which participants identify, categorize, or discriminate auditory stimuli without being required to explicitly attend to properties of whole words. (Note that similar effects are found in tasks that do require attention to whole words; e.g. lexical decision: Berent, Everett, and Shimron 2001; word segmentation: McQueen 1998; Van der Lugt 2001.) Under some accounts (e.g. McClelland, Mirman, and Luce 2006), performance in sub-word tasks directly reflects prelexical representations; phonotactic constraints therefore directly influence sub-word task performance. In contrast, other accounts (e.g. Norris *et al.* 2000) augment the two-stage architecture above with an additional set of representations that receive input from both prelexical and lexical representations. These *decision* representations support performance in sub-word level tasks (e.g. allowing participants to judge whether a phoneme was present in a stimulus). Note, however, that these representations directly receive input from prelexical representations; this account therefore also predicts phonotactic well-formedness should influence behavior in sub-word tasks.

Although sub-word tasks do not explicitly invoke lexical representations, most current accounts predict that lexical factors can exert an indirect influence on performance (this interaction between lexical and phonotactic knowledge will be discussed in greater detail below). In some theories (e.g. McClelland *et al.* 2006), lexical and prelexical processes interact with one another. Under other accounts,

sublexical decision representations receive input both from prelexical and lexical processes (e.g. Norris *et al.* 2000). Therefore, under both architectures, both lexical and sublexical properties exert an influence on sub-word task performance.

4.2.2 Evidence that Prelexical Processes Disprefer Categorically Ill-formed Structures Categorical distinctions in phonotactic well-formedness clearly exert an influence on speech perception. Many researchers have noted that relative to categorically well-formed phonological structures (e.g. attested English clusters such as /pl/), categorically ill-formed structures (e.g. unattested English clusters such as /dl/) are more difficult to perceive. This has been noted for quite some time; for example, Trubetzkoy (1939: 62–64) discusses how the constraints of one's native language impede accurate identification of sound sequences from foreign languages. Converging evidence from a variety of perceptual tasks has provided more systematic confirmation of these observations. All of these studies assume that categorical ill- vs. well-formedness of a phonological structure is roughly indexed by its absence vs. presence in lexical items of a language. For example, no English word ends in /h/; the phonotactic knowledge of English speakers therefore specifies that word-final /h/ is categorically ill-formed.

4.2.2.1 Identification Tasks The seminal study of Brown and Hildum (1956) found English listeners made more transcription errors on categorically ill-formed sequences than they did on well-formed sequences. More recent studies in other languages have confirmed this result, both in terms of accuracy (e.g. French: Hallé *et al.* 1998; Japanese: Dupoux *et al.* 1999) as well as reaction times (i.e. in phoneme monitoring tasks: Segui, Frauenfelder, and Hallé 2001). Furthermore, errors on these categorically ill-formed sequences are not random; they tend to result in categorically well-formed sequences (e.g. French speakers mistranscribe ill-formed /dl/ as well-formed /gl/: Hallé *et al.* 1998). This suggests that ill-formed stimuli activate well-formed prelexical representations.

Identification of perceptually ambiguous stimuli is biased towards categorically well-formed sequences (e.g. English: Massaro and Cohen 1983; Japanese: Dupoux *et al.* 2001). For example, when /r/ but not /l/ forms a categorically well-formed cluster (e.g. /tri/ vs. */tli/), English listeners' categorization of stimuli on a synthesized /r/—/l/ continuum is biased towards /r/. The opposite pattern is found when the well-formed cluster is composed of /l/ (e.g. /sli/ vs. */sri/). (See Moreton 2002 for recent related results.) Similar results have been found in dichotic fusion. When conflicting stimuli are presented to each ear, their perceptual fusion into a single stimulus tends to result in categorically well-formed sequences (e.g. in English, /b/+/l/ is resolved as /bl/ not /lb/: Day 1968, 1976; for similar results in Portuguese, see Morais *et al.* 1987).

4.2.2.2 Discrimination Tasks Participants have difficulty discriminating (at least some) ill-formed from well-formed sequences. Dupoux *et al.* (1999) documented Japanese speakers' difficulty (slower reaction times, greater errors) in discriminating categorically ill-formed consonant-consonant sequences from the corresponding

categorically well-formed consonant-vowel-consonant sequence (e.g. *ebzo* is difficult to discriminate from *ebuzo*). Davidson (2007) reported similar perceptual difficulties for English speakers. Kabak and Idsardi (2007) found Korean speakers have difficulties in discrimination tasks with sequences that are categorically ill-formed with respect to syllable-conditioned patterns (e.g. */g/ in coda).[6] For Japanese speakers, Dehaene-Lambertz, Dupoux, and Gout (2000) reported similar results using an implicit electrophysiological measure of discrimination; furthermore, Jacquemot *et al.* (2003) found distinct patterns of neural activity for categorically well- vs. ill-formed sequences during discrimination tasks.

4.2.2.3 Perceptual Difficulties with Categorically Ill-formed Sequences Arise in Prelexical Processing To establish the evidential status of the observations from identification and discrimination tasks, it is crucial to show the effects derive from prelexical processes. With respect to the involvement of more basic auditory processing, it is unlikely that these identification results reflect inherent uncertainties in the acoustic signal. Listener groups with varying language backgrounds exhibit different patterns of performance (e.g. Dupoux *et al.* 1999, showed that French listeners have no difficulty perceiving sequences that are difficult for Japanese participants). Acoustically identical ambiguous tokens (e.g. stimuli on an /r/—/l/ continuum) are processed differently depending on the context in which they appear (Massaro and Cohen 1983), again suggesting that it is not intrinsic properties of the acoustic signal that gives rise to these effects. Finally, with respect to lexical effects, Dupoux, Pallier, and Kakehi (2001) found comparable patterns of transcription errors regardless of whether the error results in a word or non-word – suggesting these phonotactic effects are not a simple reflection of a bias to report existing lexical items (see Day 1968, for similar results in dichotic fusion tasks).

4.2.2.4 Implications for Theories of Phonotactic Knowledge Converging results from a variety of perceptual tasks and languages provide support for the functional architecture outlined above. They are broadly consistent with the claim that categorical distinctions in phonotactic well-formedness are encoded by prelexical processes in perception. Beyond simply validating our linking assumptions, they also provide insight into the relationship between cognitive functions reflecting grammatical (e.g. phonotactic well-formedness) and lexical (e.g. word-specific) knowledge. The architecture above assumes that these are encoded by two distinct functions (prelexical and lexical); consistent with this, empirical studies show that phonotactic well-formedness exerts an influence on perceptual performance independent of lexical knowledge. This supports the claim that lexicon and grammar form distinct components of our phonological knowledge.

4.2.3 Evidence that Prelexical Processes Disprefer Relatively Ill-formed Structures

4.2.3.1 Converging Evidence for Perceptual Sensitivity to Gradient Variations in Well-formedness More recent research suggests that degrees of well-formedness within

attested sequences of a language lead to relative ease of perceptual processing. These studies have indexed degrees of well-formedness using *phonotactic probability* – the probability of a given phonological structure within a language. This can be indexed by various n-gram measures (e.g. relative probability of single phones such as /s/ or biphones such as /st/, estimated using either a corpus of utterances or a lexical database). Pitt and Samuel (1995) argued that stimuli composed of phonological structures with high phonotactic probability are detected more quickly in monitoring tasks. Pitt and McQueen (1998) reported that identification of ambiguous sounds is biased towards structures with higher phonotactic probability. Vitevitch and Luce (1999) examined English speakers' speeded same/different judgments. For nonwords, the reaction time for correct "same" responses is faster for syllables with high vs. low phonotactic probability (see Vitevitch *et al.* 2002, for similar results for cochlear implant patients with good word recognition abilities). Finally, Coetzee (2008) found that in the identification of perceptually ambiguous sounds English listeners show a dispreference for attested forms that violate phonotactic constraints on consonant cooccurrence (i.e. the Obligatory Contour Principle).

These effects can be plausibly attributed to prelexical processing. Research by Vitevitch and colleagues suggest that effects in discrimination tasks cannot be attributed to more basic auditory or lexical processes. Two critical observations argue for this point. First, behavioral effects change as a function of task. For tasks that emphasize lexical properties (e.g. lexical decisions), performance primarily reflects similarity to lexical items – in contrast to the phonotactic probability effects observed in discrimination tasks. This is found even when identical stimuli are used in both tasks (Vitevitch and Luce 1999). Second, within the same task and items, the degree to which filler items are composed of real lexical items leads to shifts in the influence of lexical vs. phonotactic factors (i.e. the use of mostly real word fillers lead to a greater influence of lexical factors; Vitevitch 2003). Since identical stimuli yield contrasting behavioral effects across tasks, it is unlikely that the effects derive from purely acoustic properties of the stimuli. With respect to lexical effects, the ability of participants to shift from phonotactically- to lexically-driven behavior within and across tasks is consistent with independent contributions of both prelexical and lexical representations to discrimination.

Distinct electrophysiological responses have been documented for high vs. low probability sequences (Dutch: Bonte *et al.* 2005). However, it is unclear if these reflect prelexical processes. Although Pylkkänen, Stringfellow, and Marantz (2002) reported that one electrophysiological measure is influenced only by phonotactic probability, Stockall, Stringfellow, and Marantz (2002) found that the same measure is also sensitive to similarity to existing lexical items.

Inconsistent results with respect to the role of lexical vs. prelexical processes have also been reported in studies of identification of ambiguous stimuli. Newman, Sawsmith, and Luce (1999) reported that lexical effects on identification of perceptually ambiguous stimuli are suppressed for stimuli with very high phonotactic probability. Pitt and McQueen (1998) reported phonotactic probability effects on identification when similarity to existing lexical items is equated. These two

studies suggest an independent contribution of phonotactic probability on perceptual processing. However, Magnuson *et al.* (2003a) found that when phonotactic probability effects conflict with similarity to lexical items, the latter dominates performance (see Magnuson *et al.* 2003b; McQueen 2003, for further discussion).

As made clear by these latter studies, sub-word tasks are clearly influenced by properties of both lexical and prelexical representations. However, finding such as those of Vitevitch (2003) reviewed above suggest that prelexical representations make an independent contribution to sub-word task performance. This is consistent with an influence of phonotactic well-formedness on perception independent of that of the lexicon.

A final source of evidence for perceptual sensitivity to gradient distinctions in well-formedness comes from studies examining the perception of sequences which are unattested in one's native language. As noted above, it has been proposed that grammars can distinguish degrees of well-formedness among unattested forms (e.g. Coetzee 2008; Davidson 2006b). Berent *et al.* (2007) examined English hearers' identification of clusters that are absent from English. These clusters vary in the degree to which they respect cross-linguistic markedness generalizations (e.g. /bn/, with a sonority rise, is less marked than /lb/, with a sonority fall). They found greater rates of misidentification for clusters which are cross-linguistically marked (see Berent and Lennertz 2007; Peperkamp 2007, for additional discussion of these findings; Berent *et al.* 2008 for similar results with Korean listeners; and Berent *et al.* 2009 for extension of these results to nasal-initial clusters). This is consistent with the encoding of gradient distinctions in well-formedness (as indexed by cross-linguistic markedness) among unattested forms. Unlike the other perceptual studies reviewed above, these experiments are not simply limited to categorical contrasts between relatively ill-formed and well-formed groups of items. Berent and colleagues found that accuracy rates vary across the various ill-formed clusters – suggesting perceptual processes are sensitive to degrees of well-formedness

4.2.3.2 Implications for Theories of Phonotactic Knowledge Studies of gradient well-formedness provide further evidence that lexicon and grammar form distinct components of our phonological knowledge. Additionally, they provide support for the claim that our knowledge of phonotactics can do more than distinguish categorically well- from ill-formed structures. Among both attested and unattested form, our knowledge of phonotactics encodes gradient distinctions in well-formedness.

4.3 Phonotactic Knowledge in Speech Production

4.3.1 Phonotactic Knowledge Within the Functional Architecture of Speech Production Like psycholinguistic theories of speech perception, production accounts assume the presence of whole-word lexical representations that mediate between phonological and syntactic/semantic information. In production, these

whole-word representations serve as input to cognitive processes that manipulate sound structure. Generally, theories assume two broad stages of processing (for a recent review of theoretical perspectives and supporting evidence, see Goldrick and Rapp 2007). The first (*phonological spell-out*) takes as input whole-word representations and yields relatively abstract, coarse-grained phonological information as output (e.g. unprosodified segments, unspecified for featural content). The second stage of processing (*phonetic encoding*) takes these abstract representations as input and yields (more) fully specified representations of sound structure that drive subsequent articulatory planning and execution processes.

Phonotactic constraints are assumed to exert an influence on phonetic encoding processes with the result that structures that are more phonotactically well-formed are retrieved more rapidly and accurately than ill-formed structures. For example, Wheeler and Touretzky (1997) proposed that phonetic encoding assigns segments (retrieved during lexical phonological processing) to prosodic positions via a constraint satisfaction procedure. When segments are misordered in speech errors, licensing constraints tend to block ill-formed sequences from being produced (e.g. blocking the error "too blue" → */tlu blu/). Goldrick and Larson (2008) accounted for effects of gradient distinctions in well-formedness by extending Warker and Dell's (2006) theory of phonetic encoding. Under this account, the strength of connections between lexical phonological and phonetic representations varies with phonotactic probability. Since error probability is related to relative activation levels (Warker and Dell 2006), the more strongly activated high probability forms are more likely to occur as errors than low probability forms.

The sensitivity of phonetic encoding processes to phonotactic well-formedness predicts a number of effects in speech production. The discussion below focuses on tasks that do not involve a substantial perceptual component (e.g. immediate repetition/shadowing: Munson 2001; Onishi, Chambers, and Fisher 2002) or metalinguistic tasks (e.g. word blending: Treiman *et al.* 2000). However, it should be noted that similar effects are observed in these studies.

4.3.2 Evidence that Phonetic Encoding Processes Disprefer Categorically Ill-formed Structures Production difficulties for phonotactically ill-formed structures have long been recognized by researchers in linguistics (e.g. Whorf 1940). Subsequent behavioral studies have more systematically confirmed these observations. As in perception studies, categorical well-formedness distinctions are indexed by presence vs. absence in a language's lexical items. In elicited production studies, targets that are categorically ill-formed yield higher error rates than well-formed targets (Amharic and Chaha: Rose and King 2007; English: Davidson 2006a). The architecture above also predicts that when errors are made, it is likely that targets will be replaced by categorically well-formed structures (as was observed in perceptual errors: e.g. Hallé *et al.* 1998). Consistent with this, numerous transcription-based studies in various languages report that spontaneous speech errors rarely result in structures which are categorically ill-formed in a speaker's native language (e.g. English speakers rarely produced errors such as "miff" → "mih"; Arabic: Abd-El-Jawad and Abu-Salim 1987; English: Vousden, Brown, and

Harley 2000; German: MacKay 1972; Mandarin: Wan and Jaeger 1998). Although higher rates of phonotactically ill-formed productions have been observed in experimentally elicited tongue twister productions,[7] transcription analysis still reports a bias towards categorically well-formed structures (e.g. English: Butterworth and Whittaker 1980).

Complementing work on native language phonotactic constraints, recent transcription studies have examined the effect of implicitly acquired categorical distinctions in well-formedness. Dell *et al.* (2000) exposed English speakers to sound distributions that suggest some structure is categorically ill-formed within the experiment. For example, in one condition, /f/ is confined to onset, suggesting /f/ is categorically ill-formed in coda. For another set of participants, /f/ is confined to coda, suggesting /f/ is categorically ill-formed in onset. They find that virtually all speech errors result in forms that are categorically well-formed with respect to the experimental condition (e.g. in the former experimental condition, virtually all /f/ error outcomes appear in onset as in "ned"→"fed"; in the latter, /f/ errors appear in coda).

As with the perceptual data reviewed above, it is critical to establish that such effects reflect the cognitive component of interest – here, phonetic encoding processes. Dell *et al.*'s (2000) results strongly support this. The fact that participant behavior reflects relatively arbitrary well-formedness distinctions specific to each condition suggests that these effects do not reflect intrinsic motoric properties of ill-formed structures. With respect to lexical effects in the same paradigm, Goldrick (2004) finds the effects of implicitly acquired well-formedness distinctions are not eliminated by the exclusion of word error outcomes. Participants' error outcomes still respected experiment-specific well-formedness when responses like "fed" were excluded from the analysis (leaving only nonwords like "fep"). This suggests a simple lexical bias cannot account for the observed error patterns. (Furthermore, a simple transcriber bias cannot account for contrasting patterns across conditions; in both studies, the same transcribers analyzed each condition.)

4.3.3 Evidence that Phonetic Encoding Processes Disprefer Relatively Ill-formed Structures Among attested sequences, relatively well-formed structures are more accurately and rapidly processed than ill-formed structures. As in studies of speech perception, this has been indexed via phonotactic probability. Vitevitch, Armbrüster, and Chu (2004) found pictures with high probability names have shorter naming latencies in English than pictures with low probability names. Laganaro and Alario (2006) found that pictures whose names are composed of high vs. low frequency syllables have shorter naming latencies in French. With respect to accuracy, elicited production studies (e.g. English: Kupin 1982) and studies of individuals with aphasia (e.g. English: Blumstein 1973) report greater accuracy for high vs. low probability structures.

With respect to error outcomes, the evidence is somewhat mixed. Some speech error analyses find errors are more likely to result in high vs. low probability sequences (e.g. English: Levitt and Healy 1985; Motley and Baars 1975). Similar results have been found in cases of aphasia (English: Buchwald 2009; Buchwald,

Rapp, and Stone 2007; Goldrick and Rapp 2007; French: Béland and Paradis 1997; Italian: Romani and Calabrese 1998). In a recent study utilizing the Dell *et al.* (2000) paradigm discussed above, Goldrick and Larson (2008) found that the probability of error outcomes is influenced by phonotactic probability. However, many studies have failed to find any asymmetries in the probability of contrasting error outcomes (e.g. Arabic: Abd-El-Jawad and Abu-Salim 1987; English: Shattuck-Hufnagel and Klatt 1979; Swedish: Söderpalm 1979). Finally, in a series of studies, Stemberger argued that although there is a general preference for higher frequency error outcomes, there is also a preference for lower frequency forms over (very high frequency) default structures (see Stemberger 2004, for a recent review).

When phonotactic probability effects have been observed, they arguably reflect the operation of the cognitive component of interest – phonetic encoding. With respect to lexical effects, Vitevitch *et al.* (2003) reported phonotactic probability effects on latency even when items are matched in terms of similarity to existing lexical items. With respect to lower level articulatory processes, at least two studies have shown that performance on the same structures varies across task conditions (Goldrick and Larson 2008; Laganaro and Alario 2006) suggesting that intrinsic articulatory properties cannot account for these behavioral effects. Finally, two additional results provide positive evidence for a phonetic encoding locus for phonotactic probability effects. Laganaro and Alario (2006) found that effects of syllable frequency are eliminated in delayed picture naming, presumably because this allows sufficient time for all speech planning processes to be completed prior to articulation. However, when a secondary task disrupted phonetic encoding, syllable frequency effects were found even in delayed naming. This suggests that phonetic encoding processes are responsible for producing the behavioral effect. Goldrick and Rapp (2007) contrasted the performance of two aphasic individuals with deficits to phonological spell-out processes vs. phonetic encoding; effects of phonotactic probability are found only when errors arose subsequent to a phonetic encoding deficit. Thus, although both lexical properties and phonotactics well-formedness influence speech production, there is ample evidence that phonotactic well-formedness independently influences speech production.

4.3.4 Implications for Theories of Phonotactic Knowledge Results from speech production converge with those from speech perception. There is clear evidence that phonotactic and lexical knowledge are encoded by distinct cognitive functions – supporting the claim that grammatical and lexical knowledge form distinct components of our knowledge of phonological structure. Several studies also suggest that phonotactic knowledge encodes gradient distinctions in well-formedness. Specifically, a number of studies show not only differences between groups of relatively well- vs. ill-formed structures, but varying degrees of performance related to degrees of well-formedness. Davidson (2006a) reported that English speakers produce distinct error rates for different clusters that are unattested in English. (She attributes these varying error rates to varying degrees of well-formedness reflecting the articulatory and perceptual properties of the

targets; see Davidson 2006b, for further discussion.) In an elicited speech error task in two Ethiopian Semitic languages (Amharic and Chaha), Rose and King (2007) found distinct error rates for forms with different degrees of OCP violations. Finally, Goldrick and Larson (2008) found that the probability of speech errors varied according to experiment-specific phonotactic probability.

4.4 Phonotactic Knowledge in Other Behavioral Domains

4.4.1 Memory Current theories postulate three core functional components for memory tasks (for recent reviews see Baddeley 2003; Vallar 2006). A *phonological analysis* component takes acoustic input of the to-be-remembered stimuli and outputs a phonological representation (similar to prelexical processing in theories of speech perception). The phonological representation output by this system serves as input to the *phonological short-term store*, a limited capacity memory system. This interfaces with a *phonological output buffer*. This maintains the activation levels of phonological representations within the short-term store and supports recall by interfacing between the short-term store and other processes supporting spoken production (for discussion of how the buffer interacts with other production processes, see Goldrick and Rapp 2007; Jacquemot and Scott 2006). Many theories assume a variety of other speech perception and production processes also contribute to memory processing (e.g. lexical processes involved in perception and production; see Martin and Gupta 2004, for a recent review).

The wide variety of processes involved in short-term memory tasks creates several opportunities for phonotactic well-formedness to influence processing. Focusing specifically on short-term recall tasks, Gathercole *et al.* (1999) outlined two broad hypotheses to account for such effects. Since phonotactic well-formedness influences prelexical perceptual processes, less well-formed structures may be less robustly stored than more well-formed structures. Alternatively, phonotactic well-formedness effects could arise in the context of either rehearsal or recall, as information from other (perceptual or production) processes sensitive to phonotactic well-formedness is used to restore or reconstruct decaying memory representations (a process termed "redintegration"; Schweickert 1993).

Many studies have shown that performance in memory tasks is correlated with measures of phonotactic well-formedness. Lists composed of structures with relatively high vs. low phonotactic probability have higher recall accuracy (English: Nimmo and Roodenrys 2002; French: Majerus and Van der Linden 2003; experiment-specific probability: Majerus *et al.* 2004). Participants show better recognition memory for English nonwords with relatively high vs. low phonotactic probability (Frisch, Large, and Pisoni 2000). Finally, high probability items have a greater likelihood of appearing intact in short-term memory errors in English and Korean (Lee and Goldrick 2008). These effects are not likely to derive from lexical factors, as phonotactic probability effects on accuracy are found when similarity to existing lexical items is controlled (although both influence accuracy: Thorn and Frankish 2005; Lee and Goldrick 2008; but see Roodenrys and Hinton 2002).

These findings provide further evidence for both the independence of grammatical and lexical knowledge, as well as the representation of gradient distinctions in well-formedness. However, the precise functional locus (or loci) of these effects are unclear – they may derive from stimulus coding and/or redintegration processes (see Baddeley 2003; Thorn and Frankish 2005, for discussion). The evidential status of these results is thus less clear than studies of production and perception, where behavioral effects could be more clearly linked to a specific cognitive function.

4.4.2 Naturalistic Corpora Ultimately the goal of a psychological theory of real-time linguistic behavior is to move beyond highly controlled experimental tasks (such as those discussed above) to fully natural communication contexts. Unfortunately, given our limited understanding of many of the cognitive functions involved in natural communication, it is extremely difficult to clearly establish the evidential status of such data. For example, differences in the relative frequency with which different phonological structures appear in corpora may not only reflect the properties of a wide array of phonological processes (involved in perception, production, or memory processing) but also a range of *non*–phonological processes[8] (involving other dimensions of linguistic structure as well as more general cognitive processes). More generally, it is critical to note that psychological realism is an issue for those that would commit sins of commission as well as omission. Just as it is incorrect to assume that some type of data is categorically irrelevant (e.g. experimental data from perception/production has no bearing on linguistic theory), it is also incorrect to assume all behavioral observations are relevant for evaluating theoretical claims.

4.5 Utilizing Psychological Realism: Progress and Prospects

Adopting the perspective of psychological realism has allowed us to draw principled inferences concerning phonotactic knowledge from speech perception and production data. In contrast to word-likeness/acceptability judgments, in these behavioral domains the functional architecture of the core processes encoding phonotactic knowledge has been articulated. Although not reviewed in detail above, the basic structure of these architectures has received broad empirical support. With respect to phonotactic knowledge specifically, the sections above reviewed data from a range of behavioral tasks that make use of these core processes. The results provide converging evidence to support the assumption that phonotactic well-formedness influences processing in both modalities. Building on this empirical support for our linking assumptions, we have used data from these tasks to address two key issues in linguistic theory. Across modalities and tasks we find broad support for two claims: the phonological grammar and lexicon are distinct components of linguistic knowledge: and our knowledge of phonotactic well-formedness can encompass gradient as well as categorical distinctions. While this represents important progress on key issues, a great many theoretical

questions are clearly still unresolved. The following sections discuss two such areas for future work.

4.5.1 Interactions Between Lexical and Grammatical Knowledge As discussed in detail in the sections above, more recent work in each of these empirical domains has established that phonotactic constraints exert an effect independent of the lexicon. In perception, effects of phonotactic well-formedness derive from prelexical, not lexical processes; similarly, in production, effects arise specifically within phonetic encoding. This is incompatible with a strong claim that phonotactic and lexical knowledge are completely functionally indistinct. However, it should also be clear from the review above that both factors contribute to performance in a variety of domains; furthermore, in many cases their contributions are not simply additive. This suggests the independent knowledge systems representing these two factors must interact during processing. An interesting area for future research will be examining the degree and nature of interaction required to account for the observed patterns of behavior across domains. In developing this line of inquiry, linguistic research should capitalize on the extensive empirical and theoretical work in psycholinguistics that has examined interaction between lexical and sublexical phonological processes (for a recent review in perception, see McClelland *et al.* 2006; production, Goldrick 2006).

4.5.2 Is Phonotactic Knowledge Distributed Across Distinct Psychological Functions? As discussed in Section 2, a critical issue in psychological realism is establishing how components of linguistic theories link up to components of psychological theories. The discussion above has identified links *within* perception and production between phonotactic knowledge and specific cognitive functions. However, the link between these functions has not been discussed. Does this work suggest that phonotactic knowledge is distributed (i.e. across distinct prelexical processes in perception and phonetic encoding processes in production)? Or, alternatively, is there a one-to-one relation between phonotactic knowledge and a single cognitive function (i.e. a sublexical process shared across perception and production)?

The evidence on this issue is decidedly mixed. The fact that phonotactic well-formedness influences both perception and production (as well as memory, well-formedness judgments, and other tasks) certainly suggests strong similarities between the capacities engaged in these various processing domains. Furthermore, across domains phonotactic constraints appear to be similarly structured (e.g. reflecting gradient as well as categorical well-formedness distinctions). However, similarity does not necessarily imply identity, making it unclear if these data support a common capacity. In favor of the specialized capacity view, there is ample evidence that distinct "repair strategies" (i.e. mappings from phonotactically ill- to well-formed structures) are invoked in perception and production (e.g. Kabak and Idsardi 2007). However, these differing strategies could reflect the differential utilization of a single function specifying well-formedness (e.g. Smolensky 1996). Existing empirical evidence is therefore unable to decide whether

the effects of phonotactic well-formedness reflect a truly general capacity or distinct but similar capacities specific to particular behavioral domains. This represents a critical area for future work exploring the relationship between concepts from linguistic theories and the psychological mechanisms involved in perception, production, and memory.

5 Discussion: The Necessity and Poverty of Psychological Realism

This chapter has argued that psychological realism should critically inform not just the methodological but also, theoretical facets of linguistic research. A psychological framework specifies the functional architecture of language processing – how linguistic knowledge is deployed in real time in specific behavioral tasks. Specifying this organization to some level of detail is a *necessary* step in drawing inferences from linguistic data (as all such data ultimately reflects real-time behavior). When this is not done, as in the case of well-formedness judgments, inferential errors are likely to occur. Stronger inferences can be drawn in behavioral domains such as speech perception and production, where core aspects of the functional architecture underlying an array of behavioral tasks have been articulated and empirically justified via converging evidence from these tasks.

Although psychological realism is a critical component of linguistic research, it should not be the only component. An exclusive focus on the cognitive capacities of individual speakers offers an incomplete account of the structure of phonological systems. For example, a number of theorists have provided explanations of phonological phenomena that explicitly reject the idea that all aspects of the linguistic system reflect the capacities of individual language users. Two broad types of such proposals have been advanced. One set attributes some systematic aspects of phonological systems to the properties of linguistic populations (for a review, see Pierrehumbert 2006). For example, Pierrehumbert (2001) attributes the coarse-grained nature of phonological patterns to their stability across individual differences in vocabulary. Another type of account focuses on the role that historical change plays in determining the synchronic structure of languages (e.g. Blevins 2004). Such contemporary accounts of language structure have deep roots within linguistic thought. Historical explanations have long played a role in explanations of language structure (see Blevins 2004: Chapter 3 for a review). With respect to the properties of linguistic populations, the tension between the role of the individual and collective has been apparent from the outset of modern linguistics. As noted by Saussure: "Language exists in each individual, yet it is common to all" (1916: 19). This leads him to speak of language both as something "psychological" which is "deposited in the brain (p. 19)" but also as an entity that is "outside the individual . . . ; it exists only by virtue of a sort of contract signed by the members of a community" (p. 14). Adopting the perspective of psychological realism should not therefore lead one to ignore the role of

linguistic *communities* in shaping language structure; our theories must strive for a balance between the group and the individual.

Claims of "psychological reality" and the use of psychological data in the development of linguistic theories have provoked strong negative reactions in both the linguistic and psychological research communities. Such reactions are at least partially attributable to a variety of misconceptions. Some are shared by practitioners of both disciplines; for example, conflating the issue of algorithmic/neural realizability with developing psychological accounts of human linguistic behavior. Other misunderstandings are primarily cross-disciplinary. On the one hand, psychologists have sometimes assumed that because linguists tend to posit generalized capacities (subserving many tasks), the constructs of linguistic theory are of a different kind than those posited by psychologists. In fact, both theories focus on highly abstract, (incomplete) characterizations of the cognitive system at the functional level of description. At most, cognitive psychological and linguistic theories merely differ in the *degree* of abstraction they adopt. This is unsurprising, as characterizing more generalized capacities may require greater abstraction. On the other hand, linguists have sometimes assumed that using psychological data requires a indiscriminate assimilation of any and all facts about behavior. In fact, cognitive psychologists are acutely aware of the difficulty of inferring cognitive structure from observed behavior.

Addressing these misunderstandings therefore serves not only to illuminate the connections between concepts in psychological and linguistic theories. It also enriches the empirical base for linguistic theories and the theoretical vocabulary of psychological theories. With respect to the case study examined here – knowledge of phonotactic well-formedness – the evidence suggests that concepts from linguistic theories are part of the psychological capacities of individuals, deployed in online behavioral tasks. Establishing such correspondences has allowed us to draw on experimental data to inform theories of phonotactic knowledge. Although not directly acknowledged above, these experimental studies have drawn (and will continue to draw) heavily on theoretical concepts from linguistics. Addressing questions of psychological realism is therefore not simply an exercise in facilitating interdisciplinary relationships. It holds enormous practical importance for the development of more comprehensive theories of the structure, acquisition, and use of language.

ACKNOWLEDGMENTS

Thanks to Ann Bradlow and especially the editors for helpful discussions. Many thanks to Lisa Davidson and Susannah Levi for helpful discussions and comments on the manuscript. Preparation of this manuscript was supported in part by National Institutes of Health grant DC 0079772.

NOTES

1 In practice, of course, theories often lie in between these two extremes. For example, some recent work in the framework of Optimality Theory has pursued the hypothesis that distinct grammars (with distinct rankings and/or constraints) are used in perception and production (e.g. Boersma 1999; Kenstowicz 2003). Similarly, psycholinguistic theories have postulated that a single system is used in both modalities (see Martin and Saffran 2002, for discussion).

2 Categorical distinctions between possible and impossible structures were also the primary focus of a good deal of pre-generative research (e.g. Bloomfield 1933; Harris 1951; Trubetzkoy 1939; Whorf 1940).

3 See Culbertson and Gross (2009) for a discussion of similar issues in grammaticality judgments.

4 These models may be generative (as described above) or based on simpler n-gram statistics (e.g. biphone frequency).

5 The potential role for multiple factors in judgments has long been recognized. Greenberg and Jenkins (1964) concluded that judgments reflect similarity to existing words. Writing about the same results a few years later, Jenkins (1966) instead appealed to "systematic relations or bodies of rules" that speakers have internalized. Chomsky and Halle (1968: 416–418) discussed a function for determining acceptability that is sensitive both to phonological rules as well as the particular structure of the lexicon.

6 Note that Kabak and Idsardi (2007) failed to find similar discrimination deficits for heterosyllabic sequences that are categorically ill-formed in Korean (e.g. *[k.m]).

7 At a subsegmental level, recent instrumental studies suggest that speech errors can result in ill-formed combinations of articulatory gestures (e.g. simultaneous tongue tip and tongue dorsum raising during production of a stop; Goldstein *et al.* 2007; McMillan, Corley, and Lickley 2008; Pouplier 2007, 2008). Findings such as these underscore the importance of using more quantitative measures of participants' behavior – not only in well-formedness judgments but in more prototypical "laboratory" studies.

8 Newmeyer (2003) raises similar points, although he draws quite different conclusions.

20 Learning and Learnability in Phonology

ADAM ALBRIGHT AND BRUCE HAYES

1 Chapter Content

A central scientific problem in phonology is how children rapidly and accurately acquire the intricate structures and patterns seen in the phonology of their native language. The solution to this problem lies in part in the discovery of the right formal theory of phonology, but another crucial element is the development of theories of learning, often in the form of machine-implemented models that attempt to mimic human ability. This chapter is a survey of work in this area. We focus here on formal approaches to modeling the path by which children learn phonological grammars of natural languages, and leave aside other important areas of research, including phonological learning from the point of view of rule induction for natural language processing, as well as mathematically characterizing the challenge of learning grammars based on finite data (Gold 1967; Angluin 1982).

2 Defining the Problem

Before we can develop a theory of how children learn phonological systems, we must first characterize the knowledge that is to be acquired. Analyses in the framework of generative phonology have traditionally focused on describing the set of attested words, developing rules or constraints that distinguish sequences that occur from those that do not. Although such analyses have proven extremely valuable in developing a set of theoretical tools for capturing phonologically relevant distinctions, it is risky to assume that human learners internalize every pattern that can be described by the theory. Indeed, it is entirely possible that there are systematic patterns that hold true of the lexicon either by sheer accident or

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

because of a series of independent historical changes (Ohala 1981; Bybee 2001; Blevins 2004; Blevins and Garrett 2004; Yu 2004). A theory of human learning should be held accountable for only that knowledge that native speakers can also be shown to have learned; cf. Becker, Nevins, and Ketrez 2008, Labov *et al*. 2006). Accordingly, we think it is best to begin by sticking to observables, that is, behaviors and intuitive judgments that reflect phonological knowledge that speakers demonstrably possess. We believe that one of the most powerful demonstrations of phonological knowledge is generalization of the pattern to unknown words. Using this criterion, we find support in the literature for at least three distinct types of phonological knowledge.[1]

Speakers possess *phonotactic knowledge*, meaning that they know, at least tacitly, what constitutes a legal word in their language. Halle (1978) gave an oft-cited example, pointing out that *brick* [bɹɪk] is an existing word of English; *blick* [blɪk] does not exist but in principle could be a word of English, while **bnick* [bnɪk] could not. Such claims can be validated not only with observations about the English lexicon, but also by observing loanword adaptation (e.g. *B'nai B'rith* [bəneɪ bɹɪθ], with [ə] inserted in /bn/ but not /bɹ/) and by observing experimentally elicited repetitions and ratings of nonce words. Such experiments have been carried out extensively on English, with Greenberg and Jenkins (1964), Scholes (1966), Pertz and Bever (1975), Ohala and Ohala (1986), and much later study. Moreover, experiments have gathered phonotactic judgments on a variety of other languages, documenting systematic cross-linguistic differences in structures that are deemed acceptable; see, for example, Moreton and Amano (1999), Frisch and Zawaydeh (2001), Berent and Shimron (2003), and Shatzman and Kager (2007). In sum, the first major task of phonological learning is to determine what is phonotactically legal in the target language.

Speakers of languages also have knowledge of *phonological alternations*. When stems and affixes are combined into words, or words into sentences, their component sounds often change in systematic ways. That speakers often internalize these patterns in their grammars is demonstrated by the substantial literature in "wug testing," starting from Berko (1958), illustrating that speakers extend patterns of alternation to nonce stems that they learn in an experimental context. For instance the American English flapping alternation (/t/ → [ɾ] / V ___ V̆) is automatically extended to novel forms; examples can be found in the wug test reported in Albright and Hayes (2003), in which nonce verbs such as *drit* ['dɹɪt] were often pronounced with a flap in suffixed forms (*dritting* ['dɹɪɾɪŋ]). For other recent demonstrations of generalization of alternations to nonce words, see Zuraw (2000), Albright, Andrade, and Hayes (2001), Pierrehumbert (2002), Ernestus and Baayen (2003), and Zhang, Lai and Turnbull-Sailor (2006).

Lastly, speakers possess knowledge of *patterns of variation*; for example, in the idiolects of English that the authors speak, it is a predictable fact about any word containing /æ/ before /m/ or /n/ that it may be realized either as [e͡ə] or as [æ̃] (the latter being preferred in more formal contexts), while other dialects show considerable differences in both the contexts and rates of tensing/raising of /æ/ (Labov 1994; Roberts and Labov 1995; Labov, Ash and Boberg 2006: Chapter 14).

As the research literature in sociolinguistics demonstrates, such cases could be multiplied indefinitely (for overviews, see Eckert 2004; Hay and Drager 2007; and Pater and Coetzee, this volume). Variation can be considered yet another form of alternation (Kawahara 2002): the same word takes on different forms, but in this case alternation is conditioned by the sociolinguistic context, rather than morphological or phonological context.

With this survey in mind, we can state the scientific problem, as we see it, as follows. The goal is to "reverse-engineer" the human system, constructing a complete model that can acquire phonology exactly as people do. The model must be able learn from positive evidence, with no overt correction of its mistakes. It must learn from real-world utterances, parsing them into their component words and morphemes. It must characterize phonological well-formedness at every level (stems, words, phrases), and it must be able to synthesize novel derived, inflected, and variant forms given suitable information about the form of a stem. Its intuitions of well-formedness must match those of humans exposed to the same data; that is, its behavior under psycholinguistic testing should be closely similar. In addition, we believe that it may often be useful to compare the errors that the model makes during the course of training to those made during acquisition by human children, as an indication that the model assumes a realistic starting point and responds to data in a fashion that is similar to human learners (Smolensky 1996b; Boersma and Levelt 2000; Tessier 2006).

To reach this goal, we need both a theory of phonology (representations, rules/constraints, internal organization of the grammar), and a theory of learning. These components are closely interdependent. No learning theory can make progress unless it is also given a hypothesis space that is adequate to characterize the elements of the learned grammar. Often, learnability researchers assume a great deal of a priori knowledge from the learner, such as a universal feature set (as in Chomsky and Halle 1968) or even a complete universal set of phonological constraints (as in Prince and Smolensky 1993/2004; Tesar and Smolensky 2000; McCarthy 2002). Moreover, the study of learnability often has consequences for the theory of grammar: in a number of cases, theorists have advocated particular principles of grammar precisely because they make phonology learnable where it would not otherwise be. We return to this possibility in Section 5.

Our chapter is organized according to phenomena: phonotactics, then alternations. However, we will also see two cross-classifying themes: the theoretical tools proposed, and particular problems faced by theorists.

3 Learning Phonotactics

3.1 Evidence Concerning Acquisition

Very little is known of the mechanisms by which humans learn the phonotactics of the ambient language. However, one result seems fairly well established: that phonotactic learning is precocious, with considerable progress made well before

children can utter words. The evidence for this emerged as new experimental techniques designed to assess the knowledge of infants, such as the head-turn preference procedure (Kemler Nelson *et al.* 1995) were applied to phonology, notably by Peter Jusczyk and his colleagues. It emerged that English-learning infants of about nine months listen longer to unfamiliar English words (Jusczyk *et al.* (1993) than to (necessarily) unfamiliar words of a similar but unfamiliar foreign language (Dutch), indicating an ability at this age to distinguish languages based on the set of sounds involved. Furthermore, Friederici and Wessels (1993) have shown that when presented with nonce words that contain legal sounds but in attested vs. unattested combinations (*bref*, *murt* vs. **febr*, **rtum*), Dutch-learning 9-month-olds pay attention longer to the attested/well-formed ones. Infants even attend to *gradient* differences of well-formedness, paying longer attention to words that contain ordinary, common phoneme sequences than to words that contain legal but rare ones (Jusczyk, Luce, and Charles-Luce 1994).

Such perception studies have important consequences for the study of phonological learning that we believe are underappreciated. In particular, they show that the tradition of observing and analyzing the spoken output of children, while valuable, may provide at best a very indirect view of what the child has actually learned about the adult language. The imperfect outputs that children later produce are indeed related to adult forms in a systematic, rule-governed way, which we believe are appropriately treated as phonological grammar.[2] However, the child's own mapping from adult forms to her own surface forms arguably is *not learned at all*, but emerges spontaneously, reflecting the child's efforts to systematically simplify her target outputs to something her still-maturing articulatory apparatus can handle.[3] In sum, current evidence suggests that the learning of the phonological pattern of the adult language is mostly an early and silent process, detectible at most indirectly in the child's own speech.

We also wish to emphasize that, other than the determination that phonotactic acquisition is precocious, very little is known. The infant experiments, ingenious though they are, have largely relied on aggregations of forms, and thus have difficulty in zeroing in on particular phonotactic configurations (though for notable exceptions, see Jusczyk, Smolensky, and Allocco 2002, and Zamuner, Fikkert, and Kerkhoff 2006). More work will be needed before such studies can join hands with theoretical modeling, which as we will see is likewise confined to making very coarse empirical predictions.

3.2 OT Models of Phonological Learning

Turning to the learning models, we will take as our starting point an influential proposal made within Optimality Theory (Prince and Smolensky 1993/2004). The scenario given here was first laid out in Tesar and Smolensky (1993) and developed by these authors in a series of works, including Tesar (1995), Tesar and Smolensky (1998), and Tesar and Smolensky (2000).

In this framework, the task of the language learner is to discover a grammar that is consistent with the set of observed forms from the target (adult) language.

We suppose that a language learner has access to a representative set of input-output pairs,[4] illustrating the mapping from underlying to surface phonological representations (more on this below). For each input representation, there is exactly one winning output, as well as a set (or a way of computing a set) of losing output candidates. The target grammar is a set of constraints, ranked in such a way that higher-ranked constraints correctly eliminate losing candidates and favor the attested winning output. The constraint set is assumed to be universal and innate (part of Universal Grammar); the task of the learner is simply to discover a ranking of these constraints that is compatible with the data (provided that one exists). The hypothesis space is the set of all $k!$ possible rankings, where k is the number of constraints.

As Tesar and Smolensky point out, the comparison of winning and losing candidates is frequently informative in identifying compatible rankings. Imagine, for instance, a language that permits the marked category of voiced obstruents. We assume for concreteness that the learner is equipped with the constraints *Voiced Obstruent ("no voiced obstruents in the output") and Ident(voice) ("output consonants must not differ from input consonants in voicing"). The system receives the datum that underlying /ba/ is pronounced [ba], and (by some means not discussed here) has available to it the fact that *[pa] is a losing candidate. The pair of candidates [ba] vs. *[pa] for (assumed) underlying /ba/ is informative with respect to constraint ranking. The constraints Ident(voice) and Ident(nasal) are *winner preferrers*, since they assign fewer violations to [ba] than to [pa] or [ma].

(1) Comparison of winning and losing candidates: Ident(voice), Ident(nasal) >> *Voiced obstruent

/ba/	Ident(voice)	Ident(nasal)	*Voiced obstruent
☞ a. [ba]			*
* b. [pa]	*!		
* c. [ma]		*!	

The constraint *Voiced obstruent, on the other hand, is (here) a *loser preferrer*, favoring *[pa]. The basic insight is that an OT grammar will derive the right outputs if, for all such pairs, every loser-preferring constraint is dominated by at least one winner-preferrer.

Tesar and Smolensky propose a ranking algorithm, Recursive Constraint Demotion (RCD), that finds grammars that have this property. RCD assumes that the learner is provided with a set of constraints and a data set consisting of winner/loser pairs generated from a grammar of fully and consistently ranked constraints (i.e. no ties, variable ranking, or errors). The algorithm is simple: it starts with all constraints unranked with respect to each other, and all winner-loser pairs unexplained. At each stage, it demotes all of the constraints that prefer unexplained losing candidates so that they are outranked by constraints that prefer only winners

or are neutral. It then checks to see which winner/loser pairs have been successfully explained by virtue of having a winner-preferring constraint ranked above all loser-preferring constraints. Once a pair has been explained, it may be removed from consideration. This reduces the set of unexplained losers and (ideally) also reduces the set of loser-preferring constraints, freeing up some constraints for ranking in the subsequent stage. This process is repeated until no explained pairs remain and all constraints have been ranked into *strata* of constraints that are not crucially ranked with respect to one another. Tesar and Smolensky show that this procedure is guaranteed to find a working grammar, provided that one exists (which it must, by the assumption that the data was generated by such a grammar).

In the example in (1), IDENT(voice) and IDENT(nasal) are winner preferrers since they both favor the correct output [ba] over an incorrect competitor. Thus, the algorithm places both in the top stratum, demoting *VOICED OBSTRUENT (which prefers losers [pa] and [ma]). With this ranking in place, both competitors are successfully eliminated by high-ranking constraints, and *VOICED OBSTRUENT no longer favors any unexplained losers, so it may be ranked, and the algorithm terminates.

As Tesar and Smolensky point out (2000: Chapters 3–4) RCD is efficient, since at each step the algorithm is guided by the pattern of constraint violations directly toward the right answer. This can be contrasted with learning procedures in which the search proceeds in a less goal-directed fashion, as in the Triggering Learning Algorithm (Gibson and Wexler 1994; Niyogi and Berwick 1996; Frank and Kapur 1996). Moreover, the algorithm is entirely general; it does not depend in any way on the particular language or set of constraints, but covers any problem that can be reduced to an input-output relation and a suitable constraint set.

The fact that RCD arrives at a compatible ranking reliably and efficiently is a strength of the approach. However, this simple version of the algorithm also relies on quite a few potentially limiting assumptions. First, it simplifies the learning task by assuming that quite a few pieces of the solution are given in advance, including the input-output pairings, the set of informative losing candidates, and the set of constraints. Second, it requires that the training data be free of errors and variation, making it inappropriate for many realistic learning tasks. Finally, while it is guaranteed to find a grammar that is compatible with the given data, it has no mechanism for deciding among multiple compatible grammars; as we will see below, this often leads to unwanted predictions. We discuss these issues in turn.

First, the finding of the losing rival candidates that RCD needs is a difficult computational problem, particularly since the set of potential candidates is infinite. Fortunately, the mathematical apparatus of *finite-state machines* has made possible considerable progress here: a finite-state machine is a formal object that (provided it includes loops) can represent an infinite class of strings and compute their vector of constraint violations. A variety of work has applied finite-state machines to the problem of finding OT candidates (Ellison 1994; Eisner 1997; Albro 1998, 2005). Riggle (2004) has shown that finite-state machines are particularly useful in finding all of the "contender" candidates – that is, those candidates which could win under at least one constraint ranking, and are thus the relevant candidates to consider in motivating constraint rankings.[5]

Second, the claim that the entire constraint inventory is given to the child in advance is understandably controversial. Given the great complexity of phonology, the assumption that the full constraint set could be innate strikes many as implausible. Here, efforts have been made to simplify constraint theory, for example by arranging constraints in families (Smith 2004), or by using the language learner's self-explored phonetic knowledge to construct constraints (Hayes 1999; Steriade 2009). The idea that constraints may be learned and that they may encode preferences that are grounded in experience with real-world limitations is particularly useful in relating grammatical learning across different components of grammar, and even across different modalities (see Brentari, this volume).

Third, Recursive Constraint Demotion relies on a comparison of winner-loser candidate pairs: for a given input, one candidate wins, and the other loses, requiring certain constraints to be ranked below others. However, when a single input has more than one possible output a contradiction emerges, since on some occasions one ranking may be needed, while on other occasions the opposite ranking is necessary. It is possible to construct OT grammars that generate free variation by letting certain constraint rankings remain unspecified, and the rankings being fixed on an utterance by utterance basis (Reynolds 1994; Anttila 1997b, 1997a; Nagy and Reynolds 1997). However, current convergence proofs rest on an assumption that the data is consistent (i.e. has no variability or errors). Other approaches to free variation are covered briefly below in Section 3.4, and in greater detail in Pater and Coetzee (this volume).

3.3 The Subset Problem in Phonotactic Learning

All of the shortcomings just discussed hold for the use of RCD for phonology in general. But the particular problem of phonotactic learning is, in one sense, even harder.

When one derives outputs from inputs (e.g. surface forms from underlying forms), it is possible to limit the problem to a set of choices, and one need only discover the correct choice.[6] But phonotactics is not obviously a matter of deriving outputs from inputs; rather, the intent is to classify all the possible phonological strings as legal or illegal. In the usual instance, the language provides only positive data, informing the learner what is legal, but no negative data to indicate what is ill-formed. It is all too easy for learning algorithms to arrive at grammars that classify the observed data as legal, while failing to classify the illegal forms as such. This is an instantiation of the classic Subset Problem for language learning (see Dell 1981; Berwick 1986; Smith 2000b; Hale and Reiss 2003; and many others).

The Subset Problem manifests itself in a particular way in standard OT, where the usual approach to phonotactics appeals to the concept of the Rich Base: it is assumed that any phonological representation can be an underlying form, and that what is legal on the surface is simply whatever can be derived, under the phonology, from any underlying representation (UR). As Smolensky (1996a, b), Smith (2000) and others point out, the "tightness" of a phonological grammar will

depend on the relative ranking of its Markedness and Faithfulness constraints; in general, the lower Faithfulness is, the fewer forms will be permitted on the surface. Unfortunately, learning based on positive evidence frequently leads RCD to rank Faithfulness constraints high, leading to an insufficiently restrictive analysis.

To see how this may happen, we return to the example from (1) above. In this language, the fact that /ba/ surfaces as [ba] and not as [pa] or [ma] is taken as evidence that IDENT(voice) and IDENT(nasal) are both ranked high. Consider now the predictions of this grammar for a hypothetic input with a voiceless nasal (/m̥a/), in a language that has no voiceless nasals. We assume that the learner comes equipped with a Markedness constraint against voiceless nasals. Since by assumption the target language does not actually contain any such sounds, the learner has no reason to demote *VOICELESS NASAL. However, if the learner is restricted to positive examples such as /ba/ → [ba] and no negative examples such as /m̥a/ → [ma], there is also no overt evidence that would compel a ranking of *VOICELESS NASAL over IDENT(voice) or IDENT(nasal). Since the Faithfulness constraints never favor a loser during training, the RCD ranks them as high as possible in the grammar, incorrectly predicting that the grammar may at least sometimes (on some occasions, or for some speakers) faithfully produce surface [m̥a]. This prediction is incorrect: studies of loanword adaptation show that when speakers are presented with sounds outside their native language, they typically modify them to conform to native language phonotactics.

(2) Inability to rule out [m̥a]

/m̥a/	*VOICELESS NASAL	IDENT(voice)	IDENT(nasal)	*VOICED OBSTRUENT
☞ a. [m̥a]	*			
☞ b. [ma]		*		
☞ c. [pa]			*	

In the present case, letting the learner start out with Markedness ranked above Faithfulness (*VOICELESS NASAL >> IDENT(voice), IDENT(nasal)) would be sufficient, at least to rule out [m̥a] in favor of [ma] or [pa]. However, study of the phonotactic subset Problem in OT quickly showed that simply starting out learning with Markedness high and Faithfulness low is unlikely to work in general: with only positive data available, Faithfulness constraints are still likely to be ranked too high (Ito and Mester 2003b). Hayes (2004), Prince and Tesar (2004), and Tessier (2006a) propose to amplify RCD with heuristics that are designed actively to keep Faithfulness constraints as low as possible, throughout learning. These proposals express a number of key insights into issues that a learner may face in deciding how restrictive the final grammar should be, but it is currently difficult to evaluate them in detail because we have relatively little empirical data concerning how well human learners solve the subset Problem.

3.4 The Gradience Problem in Phonotactic Learning

Intuitions of phonotactic well-formedness obtained experimentally are characteristically gradient: hypothetical forms can sound perfect (like, for English, [kɪp]), or completely bad (like [bzɑɹʃk]), or, crucially, intermediate (e.g. [fɹɪlg], [snɔɪks], [pwɪp]).[7] If phonological analysis is to provide a complete account of native speaker intuition, it must characterize such gradient intuitions. (The only clear alternative we are aware of is to let the grammar define a sharp binary distinction and to let analogy to existing forms cover the rest. For evidence against this view, see Hayes and Wilson (2008), Albright (2009).

An important aspect of gradient phonotactic intuitions is that they are characteristically closely related to the frequencies with which particular sequences occur in the lexicon of the language; see, for instance, Coleman and Pierrehumbert (1997), Frisch, Large, and Pisoni (2000), Bailey and Hahn (2001), and Hay, Pierrehumbert, and Beckman (2003). Jusczyk, Luce, and Luce (1994) found this to be so even in the preferences of infants, who have very little experience with the ambient language. Psycholinguists have long been interested in phonotactic gradience, and in order to measure it they have characteristically used computationally simple models that can compute predicted gradient phonotactic well-formedness scores on the basis of the frequency of existing forms in the language. Among the most common are *n*-gram models, which involve chopping up the existing lexicon into sequences *n* segments long and estimating the probability (absolute or conditional) of each *n*-gram. (For an introduction to *n*-gram models, see Jurafsky and Martin 2000: 6.) By combining probabilities across all the *n*-grams in a word, a probability value for any novel form can be computed. Such probabilities are usually positively correlated with native speaker judgments gathered experimentally (Vitevitch *et al.* 1996; Bailey and Hahn 2001).

Although *n*-gram models can mimic some aspects of gradient well-formedness intuitions, they are probably not adequate as a theory of how speakers learn and represent gradient patterns, because they fail to characterize the *multidimensionality* of phonotactic patterning. Phonological research indicates that a full model would have to include not just segmental *n*-grams, but a variety of non-local factors (Heinz 2007). In vowel harmony, vowels some arbitrary distance from one another must agree in certain of their features; and similar patterns are found for anteriority in coronal sibilants and stops, as well as laryngeal features (MacEachern 1999) (for an overview, see Hansson 2001, and Rose and Walker, this volume). Prosodic elements, like stress and tone, are also part of phonotactics, and they are characteristically non-local, requiring evaluation over windows larger than a fixed *n* segments. The question of how to integrate these multiple types of conditioning contexts in a *n*-gram model remains, as far as we know, unresolved. Furthermore, *n*-grams stated over segments suffer from insufficient generality, since they fail to incorporate features and natural classes. The practical effect is that in any language, a number of *n*-grams judged well-formed by native speakers would have zero frequency in the lexicon.[8]

A number of researchers studying non-local gradient patterns have proposed models that are more sophisticated than *n*-gram counts over sequences of adjacent segments. These include models that parse and count onsets and rhymes (Coleman and Pierrehumbert 1997; Frisch, Large, and Pisoni 2000; Treiman *et al.* 2000) or track non-local similar-segment pairs (Frisch, Pierrehumbert, and Broe 2004; Coetzee and Pater 2008). Although these models perform well at the specific tasks at hand by zeroing in on some particular aspect of phonological structure, they do not represent general purpose learning models of how speakers decide which non-local features to attend to in determining the well-formedness of a sequence. Our judgment is that models of the kind just discussed are heuristically useful, since they can represent the gradience seen in lexical patterns and which impinges on gradient intuitions, but that a full-scale theory of how learners discover phonologically significant gradient patterns will need to navigate a rich hypothesis space including representations in terms of features and natural classes, tier structure, and metrical structure.[9] Under this view, the solution will ultimately require combining insights from phonological theory about the correct representation of linguistic structure, and from machine learning about representing probabilistic information in statistical learning.

3.5 *Constraint-based Approaches to Restrictiveness and Gradience*

One effort to add gradience to constraint-based theory is a stochastic version of Optimality Theory invented by Boersma (1997, 1998). In standard OT, ranking is a purely relative notion (one constraint categorically outranks another). Boersma (1997) proposes a modification that makes ranking gradient by assigning a real number to each constraint, its *ranking value*. The grammar is made to behave stochastically as follows: on each speaking occasion, a random *noise* value, sampled from a Gaussian distribution, is added to each constraint's ranking value. The constraints are sorted according to these perturbed values, and the output is then determined according to the standard evaluation procedure for OT. Whenever conflicting constraints have sufficiently close ranking values, this system will generate multiple outputs, since the stochastic noise will cause the two constraints to switch positions on some occasions. Moreover, since ranking values are continuous, the model can produce outputs with a continuous range of probabilities. Thus at first blush, stochastic OT seems a promising candidate for solving the phonotactic gradience problem.

Stochastic OT has been used in a variety of phonological analyses as the basis for treating free variation in input-to-output mappings (see, e.g. Boersma and Hayes 2001). Moreover, just as with non-stochastic OT, stochastic OT has attracted efforts to construct learning algorithms (Boersma 1997; Lin 2005a; Maslova 2005; Wilson 2006). The Gradual Learning Algorithm (GLA; Boersma 1997) works rather similarly to Recursive Constraint Demotion,[10] only gradiently, and in a number of cases is able to learn grammars that generate free variation. Moreover, the GLA is sensitive to the frequency patterns in the learning data, which as mentioned

above are an important source of gradient intuitions in phonology. Unfortunately, unlike RCD, the GLA (at least in its current form) is demonstrably unable to arrive at the correct grammar for certain configurations of constraint violations (Pater 2008a).[11] This drawback has led researchers to seek alternative approaches, some of which constitute more radical departures from the standard assumptions of Optimality Theory.

One important idea that has recently gained attention from linguists is the principle of Maximum Likelihood (Fisher 1922), which states that the grammar to be sought is the one that maximizes the predicted probability of the learning data (also known as training data), given the constraints or other grammatical principles available. This probability is in principle a computable value under any model in which the assessed score of any form is expressed as a probability (i.e. its probability of occurring as a word). The intuitive idea is that if the probability of the observed data is maximized, then the probability of the unobserved data – or more precisely, unobserved data that can be excluded by the constraint system; cf. [blɪk] – is minimized, which corresponds to the ordinary goal of phonotactic analysis.

Jarosz (2006, 2007) adapts a Maximum Likelihood principle to OT learning. The learner starts with a provisional "pseudo-lexicon" consisting of (or sampled from) the rich base, defining a space of potential underlying forms from which any given surface form could be derived. To this is added an innate constraint set (of size k) and a body of learning data. The basis of Jarosz's approach is to search all $k!$ possible rankings of the constraints, distributing probability among them in a way that maximizes the probability of the learning data. (A very similar proposal can be found in Riggle 2006.) Moreover, since Jarosz's model assigns a probability distribution over grammars, it is able to assign gradient well-formedness predictions in the form of a probability value for each possible word. This is a highly principled solution to the phonotactic learning problem, but it comes at the cost of searching a truly colossal logical space: the set of $k!$ constraint rankings, multiplied by the space of all possible underlying forms. In order to scale up to learning scenarios with more realistic numbers of constraints and underlying forms, a more sophisticated strategy is needed for estimating probability distributions over underlying forms (most likely, making use of non-exhaustive sampling), along with some way of evaluating the likelihood of entire sets of rankings, rather than each of $k!$ rankings individually.

Hayes and Wilson (2008) propose to abandon OT altogether, adopting instead a stochastic constraint-based framework similar to Harmonic Grammar (Legendre, Miyata, and Smolensky 1990; Smolensky and Legendre 2006), in which constraints are weighted rather than given OT rankings. Specifically, Hayes and Wilson employ a Maximum Entropy (log-linear) model to find weights for inductively learned Markedness constraints that evaluate the probability of surface strings (as opposed to evaluating them as outputs for some hypothesized input, as in standard OT). A benefit of adopting weighted rather than ranked constraints is that standard search algorithms exist that provably converge on the best-fitting set of weights; for discussion, see Goldwater and Johnson (2003), Jäger (2004), Pater, Bhatt, and

Potts (2007), Hayes and Wilson (2008), and Boersma and Pater (2008). It is worth noting that in this case, the choice of weighted constraints rather than variably but strictly ranked constraints to handle variable or gradient patterns is motivated almost entirely by convergence properties (that is, convergence with gradient data) rather than because Optimality Theoretic grammars have been shown to be inadequate for the task at hand. This appears to be one of the first instances in which considerations of learnability have played a role in motivating architectural decisions in phonological theory.

4 Phonological Alternations

The learner can make significant progress on the task of learning surface phonotactics by applying the techniques described above to a set of training data consisting of individual words, or perhaps even a rougher parse of the speech stream into (approximately) word-sized units. However, phonotactic learning is not the only task that learners face: they must at the same time refine their segmentations to determine which words are morphologically complex,[12] and begin to compare related words to discover contextual variation in their pronunciation. For example, a child acquiring Dutch would discover that the word-final [t] in [bɛt] 'bed.SG.' corresponds to [d] in the suffixed form [bɛdən] 'bed.PL.'. Discovering and encoding alternations such as [t] ~ [d] is, logically speaking, a more complex task than learning static phonotactics, since it requires comparing morphologically related forms, choosing a basic or underlying form, and learning a grammar that can generate the various surface realizations.

In many cases, prior knowledge of phonotactics could give the learner a leg up in discovering alternations, since as has long been noted, alternations frequently find transparent motivation in phonotactic considerations (Kisseberth 1970; Sommerstein 1974). For example, the Dutch voicing alternation seen in [bɛt] ~ [bɛdən] 'bed-SG./PL.' is straightforwardly related to a very general ban on final voiced obstruents in Dutch (*[bɛd]). Optimality Theory provides a straightforward way of relating phonotactic learning with learning of alternations, since an initial phase of phonotactic learning can provide the learner with a crucial component of the analysis (*FINAL VOICED OBSTRUENT >> Faithfulness); all that remains is to learn the relative ranking among faithfulness constraints. For example, a child learning Dutch would need to learn that final voiced obstruents are fixed by devoicing rather than, say, nasalization ([bɛ**n**] ~ [bɛdən]) or vocalization ([bɛ**j**] ~ [bɛdən]). This would follow from the rankings IDENT(nasal), IDENT(consonantal) >> IDENT(voice).

There is reason to believe that children take on the task of learning alternations only after they have made a certain amount of headway on learning phonotactics (Hayes 2004; Prince and Tesar 2004; Tesar and Prince 2007). As discussed above, infants show sensitivity to native language phonotactics at ages as young as 9 months, well before they demonstrate systematic knowledge of words or morphological paradigms. Additional evidence that phonotactic distributions are

mastered prior to alternations comes from early child productions. Berko (1958) tested the ability of English-learning 4-year-olds to apply voicing and epenthesis alternations in the plural and past tense inflections of novel nouns and verbs: *spow+ed* [spoʊ**d**], *rick+ed* [rɪk**t**], bodd+ed [badə**d**]. For the most part, children's responses either applied the alternations correctly or omitted the suffix completely; that is, children consistently obey the phonotactics of the adult language (voicing agreement in final obstruent clusters, a ban on identical adjacent consonants), even at a stage when they have not completely mastered the alternations. Furthermore, experimental work with adult speakers has shown that prior phonotactic knowledge (in this case, from the native language) facilitates learning of alternations in an artificial language (Pater and Tessier 2003, 2006).

In the following sections, we briefly review some evidence concerning the acquisition of alternations in children, before turning to proposals for how to model the learning of alternations.

4.1 Evidence Concerning Acquisition of Alternations

Compared with knowledge of surface phonotactics, which can be demonstrated in early infancy (see above), relatively less is known about early knowledge of phonological alternations. By looking at child productions, it is possible to show that at least some alternations are acquired fairly early. For example, Aksu-Koç and Slobin (1985) describe a Turkish-learning 15-month-old who shows correct mastery of vowel harmony in the accusative suffix ([-a] vs. [-e]). However, wug tests investigating productive mastery of alternations often reveal errors even when children are correctly deploying variants of existing words. It appears that adult-like mastery of many alternations does not emerge until significantly later, with children initially preferring invariant (non-alternating) morphemes. Zamuner, Kerkhoff, and Fikkert (2006) and Kerkhoff (2007) have shown that Dutch-learning children have difficulty both recognizing and applying final devoicing in the singulars of novel nouns, while they perform much better on non-alternating items. Kazazis (1969) presents a case study of one Greek-learning child who at age 4;7 systematically failed to apply the phonotactically regular alternation between [ç] before front vowels ~ [x] elsewhere, resulting in erroneous forms such as *é*[x]*ete* 'have.2PL.' instead of adult *é*[ç]*ete*. In both of these cases, there is reason to believe that knowledge of the relevant alternations is acquired eventually; for instance, palatalization alternations are completely predictable and are applied automatically by adult Greek speakers.

Lexically restricted alternations, which frequently have no synchronic phonotactic motivation, appear to pose an even greater difficulty. Clahsen, Aveledo, and Roca (2002) show that Spanish-learning children often fail to apply irregular changes such as diphthongization (stressless [e], [o] ~ stressed [jé], [wé]) within verbal paradigms, and Clahsen *et al.* (2002a) demonstrate that German learners are likewise reluctant to apply umlaut alternations (e.g. [a] ~ [e]) within present tense verbal paradigms. Similarly, Berko (1958) found that English-learning 4 year olds were relatively unlikely to produce voicing alternations in stem-final fricatives in the

plural of novel nouns (*heaf ~ heaves*) (see also Baker and Derwing 1982; Derwing and Baker 1986). This can be contrasted with adult speakers, who do at least sometimes extend lexically restricted alternations in similar experimental settings (Berko 1958; Zuraw 2000; Albright, Andrade, and Hayes 2001; Albright and Hayes 2003; Pierrehumbert 2006). Thus, it appears that the knowledge of such alternations is acquired much later than knowledge of phonotactics and phonotactically motivated alternations.

The picture that emerges is that the task of learning alternations is a difficult one that requires significant lexical knowledge, and which is taken on gradually over the first 5–10 years of life. Furthermore, although prior knowledge of phonotactics is certainly helpful, it by no means predetermines knowledge of alternations. It appears that even when the relevant phonotactic is known, learners must nonetheless compare related forms and encode alternating variants in some fashion. The procedure that is needed to do this depends intimately on the grammatical mechanism that is employed to encode alternations.

In this section, we review two major approaches, and some challenges.

4.2 *Theories for Learning Alternations I: Approaches Using Underlying Forms*

One very widely used strategy for encoding alternations is to provide each morpheme with a single unified representation (the underlying form/representation, or UR), and to set up a grammar that derives all observed surface variants from the same underlying form. (Pāṇini; Bloomfield 1933; Chomsky and Halle 1968). For example, a learner of Dutch confronted with related forms [bɛt] ~ [bɛd-ən] 'bed.SG./PL.' would be forced to select a single underlying form – /bɛt/, /bɛd/, or something more abstract – and derive the surface alternation by intervocalic voicing (/bɛt-ən/ → [bɛdən]) or final devoicing (/bɛd/ → [bɛt]). In the Dutch case, the choice can be made relatively straightforwardly by observing the simultaneous existence of non-alternating voiceless morphemes ([vuːt] ~ [vuːtən] 'foot-SG./PL.'), making intervocalic voicing an untenable solution. In the general case, however, learning a suitable combination of URs + grammar can be difficult because of the circularity involved: the optimal choice for URs depends on having a reasonably good hypothesis about the grammar, but the grammar cannot be formulated without a hypothesis about the set of input → output mappings that it must perform.

The problem of simultaneously learning underlying forms and a grammar that makes use of them is an instance of the more general problem of *hidden structure*: the grammar depends on distinctions that are not part of the immediately observable phonetic context, but rather are structural entities encoded on a language-particular basis. In the Dutch case, the difference between the behavior of the final stops in [bɛt] ~ [bɛdən] and [vut] ~ [vutən] 'foot' is attributed to an underlying distinction (/t/ vs. /d/), which is not directly observable (since learners have access only to surface forms), but must be inferred from its effect on surface forms. The grammar must be set up in such a way that /t/ and /d/ are neutralized in some

contexts and distinct in others. Other instances of hidden structure that have been proposed in the literature include intermediate levels of representation in serial derivational frameworks, the assignment of segmental material into prosodic structure (feet, syllables, sub-syllabic constituents, weight-bearing units), and segmental feature specifications, including distinctions between full vs. underspecification and also language-particular assignments of phonological feature values to segments (Dresher 2004; Mielke 2005b, 2008; Rice 2005). In all of these cases, the correct grammar cannot be found until the hidden structure has been established, while hypotheses about hidden structure cannot be evaluated until the corresponding grammar is constructed. In order to break into this circularity, the learner must have some independent means of establishing hypotheses about either the grammar or the hidden structure[13] (Tesar *et al.* 2003; Apoussidou 2007).

4.2.1 Proposals for Learning Underlying Forms Underlying forms are an especially challenging type of hidden structure to recover, since in principle there are infinitely many possible hypotheses about the underlying form of any given morpheme. Two assumptions have proven useful in helping the learner break into the system: (1) the prior stage of phonotactic learning provides an initial hypothesis about key aspects of the grammar, and (2) lexicon optimization, which favors underlying forms that are as close as possible to their corresponding surface forms, provides the learner with a set of initial hypotheses about underlying forms of a morpheme (Prince and Smolensky 1993, 2004).

Proposals for establishing underlying forms typically rely on some form of the following strategy to establish initial hypotheses. First, if a morpheme never alternates in a particular feature, its underlying value is equal to its sole surface value (modulo robust interpretive parsing – that is, the assignment of predictable hidden structure to ambiguous phonetic surface forms; Tesar and Smolensky 2000, pp. 12–14). This lets the learner establish a "skeletal frame" of invariant features for each morpheme (Inkelas 1995; Tesar and Prince 2007; see also Kenstowicz and Kisseberth 1977: Chapter 1 for discussion). Second, if a morpheme does alternate in a feature, we need some way of selecting an underlying value. This requires that the learner have available a set of hypotheses about possible underlying forms, and is able to evaluate which of these will lead to a grammar that is consistent with all of the known data.

One straightforward approach is to let the learner pick a value (arbitrarily) from among the set of attested surface values, and try to learn a grammar that goes along with this assumed UR (Kenstowicz and Kisseberth 1977: 33). If the first value that is chosen creates a ranking paradox so that it is not possible to learn a consistent ranking that covers all of the data, the learner retracts the hypothesis and tries a different value (Kager 1999; Tesar *et al.* 2003; Tesar and Prince 2007). In the case of Dutch voicing neutralization, the procedure works as follows: for invariant morphemes like [vut] ~ [vutən], the UR must be /vut/. For alternating morphemes like [bɛt] ~ [bɛdən], the learner has two choices: [–voice] and [+voice]. Suppose the learner starts by hypothesizing /bɛt/ ([–voice]). In order to validate this hypothesis, the learner seeks to construct a grammar that maps /bɛt+ən/ →

[bɛdən], while at the same time mapping /vut+ən/ → [vutən]. These requirements are mutually incompatible, since the grammar must simultaneously allow intervocalic voicing (*VTV >> FAITH(voicing)) and maintain intervocalic voiceless stops (FAITH(voicing) >> *VTV). Thus, the hypothesis leads to inconsistency and can be rejected, leaving the learner to consider the hypothesis /bɛd/. In this case, it is no problem to learn a grammar that is compatible with the full set of known surface forms, since all that is required is that stops surface faithfully before sonorants and devoice elsewhere (FAITH(voicing)/_[+sonorant] >> *[+voice, –sonorant]). In fact, this is the ranking that the Dutch learner would already have from the prior stage of phonotactic learning, and a learner that makes maximal use of previous knowledge might favor this solution even if a consistent ranking could be learned with a different UR (Pater 2000; Tesar and Prince 2007). Thus, by trial and error the learner is able to arrive at a working combination of URs and grammar.

A related approach, proposed by Jarosz (2006), is to let the learner acquire lexical representations by entertaining all possible hypotheses of grammars and underlying forms simultaneously, using Maximum Likelihood Estimation to assign each combination a probability given the current set of data. In a case like Dutch, the learner considers URs with voiced and voiceless values, and grammars with final devoicing, intervocalic voicing, both processes, and neither process. As noted above, there is no combination of underlying forms that can generate the attested surface forms with intervocalic voicing or fully contrastive voicing, so the only grammar+UR combination that is assigned high probability after the model receives data from morphologically related forms is one that has an underlying voicing contrast and final devoicing.

These procedures work in the Dutch example, but they are not particularly efficient. Randomly trying out different feature values may require as many as 2^n guesses, where n = the number of alternating feature values in the lexicon, and as many runs of Recursive Constraint Demotion. Likewise, assessing probability distributions over all logically possible grammar+UR combinations is a computationally intensive task which is infeasible to carry out exhaustively in all but the simplest cases. Ultimately, the learner would benefit from a way of letting successful discovery of underlying values inform choices for other words. This can be seen most clearly in cases where multiple morphemes must have their underlying values set correctly before a consistent ranking can be found. Suppose the Dutch learner knew a number of plural forms with voicing ([bɛdən] 'beds', [hudən] 'hats', [hɑndən] 'hands', [krɑbən] 'crabs') at the time when the plural morpheme was learned, so that there are multiple alternating stems in the data. A consistent grammar cannot be found until every one of these morphemes is listed with an underlying [+voice] value. McCarthy (2005) proposes a procedure by which decisions about underlying values may be extended to multiple morphemes at once, which could help guide the learner to this hypothesis.[14] In addition, the learner might make use of the fact that voicing contrasts are already known (from the prior stage of phonotactic learning) to surface faithfully only in pre-sonorant position to favor the value found in the plural. Finally, a more efficient learner

might make use of the fact that some feature values are known never to contrast in any context on the surface, and are therefore unlikely to be useful in characterizing attested alternations (Dresher 2004).

Another weakness of these approaches is that progress is "all or nothing," since a hypothesized UR is deemed successful only once a consistent grammar is found that yields that attested surface form. This may be an overly stringent criterion in cases where morphemes participate in multiple alternations, since the learner may be able to make sense of certain aspects of the word but may not yet have sufficient data to arrive at a full analysis. For example, some Dutch nouns alternate not only in voicing, but also vowel length/quality: [bɑt] ~ [b**aː**dən] 'bath-SG./PL.', [smɪt] ~ [smedən] 'smith-SG./PL.'. It is plausible to think that learners may be able to establish the underlying voicing value even if they do not yet understand the (now lexically restricted) vowel alternation.[15] Apoussidou (2007: 167–168) proposes that knowledge of different underlying feature values of a morpheme are encoded separately, so that the learner need not arrive at a fully consistent ranking for all feature values simultaneously. The "all-or-nothing" criterion of success is also difficult to meet in cases where the choice of underlying values for one morpheme depends on the choice of values for another morpheme; in such cases, it is useful to allow the learner to focus on pairs of forms that differ by only a single morpheme at a time, in order to restrict the hypothesis space of possible modifications (Alderete *et al.* 2005; Merchant and Tesar 2008).

A particularly challenging configuration concerns cases of three-way contrast: alternating morphemes A ~ B exist alongside both non-alternating A and non-alternating B. An example is provided by Turkish (Kaisse 1986; Inkelas 1995):

(3)	sanat ~ sanat-ɯ	'art-NOM./ACC.'
	kanat ~ kanad-ɯ	'wing-NOM./ACC.'
	etyd ~ etyd-y	'etude-NOM./ACC.'

Applying the reasoning above, the presence of non-alternating [t] and non-alternating [d] would straightforwardly lead the learner to posit URs such as /sanat/, /etyd/, which then requires a grammar that allows underlying voicing values to surface faithfully in all contexts (FAITH(voice) >> *VTV, *[+voice,–sonorant] /_[–sonorant]). The challenge is to infer underlying values for morphemes with alternating [t] ~ [d]: positing [–voice] would require a process of intervocalic voicing (incorrectly ruling out [sana**t**ɯ]), while positing [+voice] would require a process of final devoicing (incorrectly ruling out [ety**d**]). Numerous solutions to such configurations have been put forward in the literature, including underspecification of alternating segments to exempt them from Faithfulness (Inkelas 1995), or listing both values as underlying for alternating morphemes (Hooper 1976; Kager 2009). Such representations have proven effective in distinguishing many cases of alternating vs. non-alternating morphemes, but they come at a cost: the search space for underlying representations goes beyond the set of surface-observable feature values to include underspecified representations or even "over-specified" representations that include floating features or other structure that

does not appear on the surface. (See Kenstowicz and Kisseberth 1977: Chapter 1 for relevant discussion.)

One final challenge that is worth mentioning are cases in which the learner may wish to consider underlying values that are distinct from surface values, even in the absence of surface alternations. Kenstowicz and Kisseberth (1977) discuss an example from Yawelmani Yokuts in which the future suffix *-en/-on* undergoes rounding harmony to match preceding high vowels (*xil-en* 'will tangle' vs. *mut-on* 'will swear'), but not non-high vowels (*bok'-en* 'will find'), contrary to the usual pattern in the language of rounding harmony among vowels that agree in height. This fact suggests that the future suffix is underlyingly high, conditioning the expected rounding harmony with high vowels and then lowering. Unfortunately, the simplest version of this hypothesis – namely, that the suffix is underlyingly /-in/ – is not tenable, since Yokuts has many short high vowels [i] and [u] that do not lower to [e], [o]. Kenstowicz and Kisseberth (following Kuroda 1967) make use of the fact that Yokuts has long vowels, and that they are subject to two restrictions: they generally do not occur in closed syllables (*eːn, *oːn), and there are no long high vowels in suffixes of this type (*iː, *uː).[16] Putting these together, they posit that the future suffix *-en/-on* has an underlying long high vowel /iːn/, which undergoes rounding harmony with preceding high vowels, and then lowers to mid and shortens due to the coda consonant (see Kenstowicz and Kisseberth (1977: 47–48) for details and arguments). This solution provides an elegant account of why the future suffix alternates in an unexpected way, but requires that learners consider underlying long vowels for morphemes that are always short on the surface. As Kenstowicz and Kisseberth point out, if such analyses are accepted, there are few (if any) criteria that can be imposed on possible divergences between underlying and surface forms. This makes it difficult to define formal procedures that can efficiently discover the full range of types of underlying forms that have been used in phonological analyses (see also Hockett 1955). A sensible heuristic (favored also by the principle of lexicon optimization) would be to favor underlying values that are as close as possible to attested surface values (Dresher 1981). Featural distance alone is not likely to be sufficient to guide the search, however, since the search space for underlying forms that differ by even a single feature value from the set of attested surface values may be quite large if floating features or abstract diacritic features are permitted.

Traditionally, considerations of learnability have not played a major role in helping to choose among possible theories of how to encode surface distinctions with underlying representations. We anticipate that as work proceeds on automated algorithmic discovery of underlying forms, the learnability ramifications of more complex representations may well be a more prominent factor in adopting one strategy over another.

4.2.2 Opacity The example of the Yokuts future suffix discussed in the previous section is difficult not only because the hypothesized underlying long vowel never surfaces, but also because the interaction with rounding harmony is *opaque* (Kiparsky 1971: 621–623): the suffix agrees in rounding with a preceding high vowel, but

surfaces as a [–high] vowel which would otherwise be exempt from [+high] rounding harmony: /t'ujt'uj-iːn/ → [t'ujt'ujon] 'will shoot repeatedly' (cf. [hud-al]/ *[hud-ol] 'might recognize'). This is an example of counter-bleeding opacity: harmony occurs even though an independent process intervenes, removing the apparent motivation for the change. At the same time, the suffix fails to agree in rounding with preceding round vowels that do match in height: /bok'-iːn/ → [bok'en]/*[bok'on] 'will find'. This is an example of counter-feeding opacity: lowering of /iː/ to [e] creates a mid vowel that would ordinarily be subject to harmony, but it fails to harmonize due to its underlying [+high] status.

(4) Opacity in the Yokuts future suffix

UR	/t'uyt'uy-iːn/	/bok'-iːn/
Rounding harmony	t'uyt'uyuːn	*n.a.*
Lowering of high vowels	t'uyt'uyoːn	bok'eːn
Closed syllable shortening	[t'uyt'uyon]	[bok'en]

In both cases, the relation between phonotactics and alternations is disrupted. In the case of counter-bleeding opacity, the learner encounters apparently unmotivated alternations that prior knowledge of surface phonotactics cannot help to explain (the context for the alternation is not surface-apparent), while in the case of counter-feeding opacity, the learner encounters surface exceptions that stand in the way of learning the alternation in the first place (i.e. it is not surface-true). The intuition has often been expressed in the literature that these features of opacity must be an obstacle to learning opaque interactions (Kenstowicz and Kisseberth 1977: 169; Hock 1991: Chapter 11).

In many cases, it is plausible to suppose that the learner is aided by a large number of forms in which just one of the two processes applies, allowing a certain amount of grammatical learning based on unambiguous data (Bermúdez-Otero 2003). For example, in Yokuts, the non-future suffix *-hin/-hun*, the perfective suffix *-mi/-mu*, the future passive suffix *-nit/-nut* and the dubitative suffix *-al/-ol* all show the general pattern of height-conditioned rounding harmony. This could conceivably strengthen the conviction of the learner that the observed alternations are in fact all motivated by the same phonological constraints, and help guide the learner to posit abstract levels of representation in which the same conditions are present for the opaque cases.

Even when simpler unambiguous cases are available, however, the task of learning opaque interactions between multiple processes is necessarily more difficult than the task of learning a single alternation. Indeed, closer scrutiny of what speakers actually extract from data involving opaque interactions may shed light on the workings of the system (Mielke, Armstrong, and Hume 2003). For example, Sanders (2003) tested the willingness of Polish speakers to generalize an opaque vowel-raising process to novel words, and found that the alternation, though amply attested in the lexicon, was not extended productively. On the other hand, Poliquin (2006: 136–143) found that Canadian French speakers readily apply

an opaque vowel harmony process to low frequency and novel words. Clearly further experimental work on the synchronic productivity of opaque processes will be an important source of evidence concerning whether (and how) speakers learn them.

Another important traditional source of evidence about what is learned comes from language change. It has long been observed that opaque interactions are unstable, and are frequently reanalyzed such that both processes apply transparently, or one of the processes is lost (Kiparsky 1965; King 1969: 87–101). Hansson and Sprouse (1999) contrast the fate of rounding harmony in a later generation of Yokuts speakers, observing that harmony among high vowels is preserved (/ʔukn-hin/ → [ʔukun-h**u**n] 'drink-NON-FUT.') while harmony among non-high vowels is lost (/woːn-k'a/ → [won-k'**a**]). The difference appears to be in how the two processes interacted with vowel lowering. As noted above, lowering counter-bleeds high vowel harmony, causing it to apply in more cases than expected, while it counter-feeds low vowel harmony and creates surface exceptions. Hansson hypothesizes (consistent with claims by Kiparsky, King, and others) that harmony among high vowels was easier to learn in the original system because it applied consistently in at least a subset of the relevant contexts. This provides another piece of evidence that bootstrapping from a subset of the data that shows transparent and reliable application may provide an important entry into the system.

One additional factor that appears to facilitate the learning (and creation) of opaque rule orderings is the fact that counter-bleeding interactions frequently reduce alternations, leading to greater Paradigm Uniformity (Kiparsky 1972; King 1973; Kenstowicz and Kisseberth 1977: 163–164; Burzio 1996; Kenstowicz 1996). McCarthy (1998) argues for independent reasons that learners must be biased to place Paradigm Uniformity (output-output faithfulness) constraints at the top of the ranking, above Markedness constraints. This correctly predicts that learners should easily be able to analyze – or may even accidentally create – opaque interactions that eliminate paradigmatic alternations, as in the Greek example described in Section 4.1.

It should also be emphasized that the fact that speakers frequently stop applying opaque processes should not be taken as evidence that learners fail to notice them entirely. In fact, there is reason to think that when learners are confronted with conflicting data caused by counter-bleeding interactions, they seek to explain the competition by exploring complex and detailed conditioning environments; we return to this issue in Section 4.5.

4.3 Theories for Learning Alternations II: Approaches Using Surface Mappings

An alternative approach to encoding alternations within paradigms is as relations among surface forms. Returning to the Dutch example [bɛt] ~ [bɛdən] 'bed-SG./PL.', one could observe that stem-final [d] in the plural corresponds with [t] in the singular (though not always the reverse), and encode this directly as a relation between surface forms. One common approach to limiting phonological processes

to relations between surface forms is to require that the underlying form match one attested surface allomorph (Harris 1942, 1951: 308, Footnote 14; McCawley 1967; Vennemann 1974; Hooper 1976; Kenstowicz and Kisseberth 1977: 28–33). In other theories, alternations are simply built in to the statement of the morphological mapping: [Xd-ən] in the plural → [Xt] in the singular (Zwicky 1985b; Wurzel 1987; Bochner 1993; Barr 1994; Albright and Hayes 2002). Alternatively, work within the framework of Optimality Theory has proposed capturing such surface relations using the machinery of output-output correspondence constraints (Burzio 1996; Russell 1999; Cole and Hualde 1998; MacBride 2004).

When underlying forms (or inputs to morphological mappings) are limited to surface forms, the search space for underlying forms is greatly reduced. This is not guaranteed to reduce the learning challenge, however, since the learner must instead find reliable implicational relations between surface forms. A learner of Dutch, for example, would need to learn that a voiceless obstruent in the plural ([vu**t**-ən] 'foot-PL.') reliably corresponds to a voiceless obstruent in the singular ([vu**t**] 'foot.SG.'), but the reverse does not hold ([bɛ**t**] ~ [bɛ**d**-ən] 'bed.SG./PL.', not *[bɛ**t**-ən]). In learning the predictors of voicing, two kinds of search are useful: a search for phonological contexts that frequently accompany voicing, and a search for those surface forms that most reliably reveal voicing.

First, learners may search for phonological contexts that are correlated with the difference in voicing between [vu**t**-ən] and [bɛ**d**-ən]. Ernestus and Baayen (2003) show that voicing of stem-final obstruents can be predicted to a significant extent based on the place and manner of the segment in question, as well as features such as the preceding vowel length. They provide experimental evidence that speakers are able to use these lexical trends to predict the probability of alternations in nonce words. (We return to the issue of lexical gradience below in Section 4.5). One procedure for discovering reliable predictors of an alternation is what Albright and Hayes (2002, 2003), building on a proposal sketched by Pinker and Prince (1988), call the *minimal generalization* approach: the learner compares pairs of morphologically related surface forms to determine what they have in common and what varies between the two forms. For example, a Dutch learner confronted with the pair [vut] ~ [vutən] 'foot-SG./PL.' would align the material in the two forms to discover that they differ only in the addition of a suffix (∅ → ən/__#), while alternating forms like [bɛt] ~ [bɛdən] 'bed-SG./PL.' differ both in voicing and the addition of a suffix (t → dən/__#). By comparing additional pairs such as [rat] ~ [radən] 'wheel-SG./PL.', the learner attempts to extract phonological features that statistically favor alternation or non-alternation. Based on just these three items ([vut] vs. [bɛt], [rat]) , the height and rounding of the preceding vowel look like they might be reliable indicators, with voicing alternations occurring after [–high] or [–round] vowels. Consideration of more data would reveal that this particular correlation turns out not to be particularly strong in the Dutch lexicon, but other features such as vowel length are strongly correlated with voicing alternations (Ernestus and Baayen 2003; Kerkhoff 2007: 96–104). Other algorithmic approaches to identifying predictive contexts include decision tree-based approaches (Breiman *et al.* 1984; Ling and Marinov 1993; Gildea and Jurafsky 1996), the

Analogical Modeling of Language (AML: Skousen 1989; Eddington 2003), TiMBL (Daelemans *et al.* 2000), and the Generalized Context Model (Nosofsky 1986, 1990; Nakisa, Plunkett, and Hahn, 1997).

A second type of information that can help ensure accurate inferences based on surface forms is the knowledge that some forms are better than others at revealing surface contrasts. For example, in Dutch nouns it is clear that the plural is a better source of information than the singular about the voicing of stem-final obstruents, since the singular undergoes final devoicing while the plural maintains voicing contrasts. Thus, a learner might wish to learn about asymmetries in the predictive power of different members of the paradigm. Albright (2002) proposes a procedure in which learners compare the reliability of mappings based on different available surface forms by using the minimal generalization algorithm to construct grammars using each part of the paradigm as an input, and evaluating the accuracy of the resulting grammars. In this way, the learner can discover that some parts of the paradigm undergo more neutralizations than others, and can subsequently focus on just those mappings that are known to have high predictive value. As has long been noted (e.g. Kenstowicz and Kissberth 1977: 28–33), theories that operate on surface allomorphs are much more restrictive than those that operate on more abstract underlying representations. A potential advantage of this restrictiveness is that it greatly simplifies the learning task, since the learner need only identify those parts of the paradigm that tend to be most informative in the language rather than comparing all forms of all words to locate contrastive values on a morpheme-by-morpheme basis.

There is reason to believe that learners do indeed focus on particular parts of the paradigm that are characteristically most informative. A particularly revealing source of evidence comes from cases of "consistent inheritance," in which idiosyncratic properties of one paradigm member are carried over to other paradigm members. Spanish provides a telling example. Many Spanish verbs show alternations between a velar stop in some forms and ∅ in others:

(5) Spanish velar alternations

salir 'to leave'	*Present indicative*	*Present subjunctive*
1SG	sal**g**-o	sal**g**-a
3SG	sal-e	sal**g**-a
1PL	sal-imos	sal**g**-amos

An approach using underlying forms might posit an underlying /g/ that deletes before front vowels: /salg-e/ → [sale], producing paradigms in which [g] is retained only before back vowels (Harris 1969). An approach based on surface mappings would instead rely on implicational relations among surface forms: the present subjunctive matches the form found in the 1SG present indicative. Although this statement misses the relation between presence of [g] and the following vowel quality, it makes a much more general prediction: the subjunctive should always resemble the 1SG. indicative. This prediction is in fact correct: in verbs where the

1SG. indicative differs from the remaining indicative forms in other idiosyncratic ways, the present subjunctive consistently inherits the properties of the 1SG indicative (Maiden 2005).

(6) Idiosyncratic alternations in Spanish

caber 'to fit'	Present indicative		Present subjunctive	
1SG	quep-o	[**kep**o]	quep-a	[**kep**a]
3SG	cab-e	[kabe]	quep-a	[**kep**a]
1PL	cab-emos	[kabemos]	quep-amos	[**kep**amos]

The phenomenon of consistent inheritance is sometimes referred to as parasitic or Priscianic derivation (Matthews 1972; Aronoff 1994). That these resemblances are not accidental is shown by the fact that speakers appear to actively enforce them, analogically replacing exceptional forms with novel ones that conform to the inherited relationship. This may be taken to indicate that speakers learn systematic relations among particular surface forms within the paradigm (Zwicky 1985b; Stump 2001).

One final type of information that may allow learners to arrive at a unified analysis of diverse surface mappings is knowledge of independent phonological processes. For example, the minimal generalization procedure described above could learn that the [d], [t], and [əd] allomorphs of the English past tense suffix reliably occur in different environments, but this fails to capture the intuition (codified in English orthography) that they are underlyingly the same morpheme. If the model was able to hypothesize that word-final obstruent clusters undergo voicing assimilation and epenthesis, it would be able to derive all regular past tense forms using a single surface mapping: $\varnothing \rightarrow$ d. Goldsmith (2006: 13–16) proposes an algorithm for collapsing the representations of similar morphemes, driven by the goal to economize on the number of distinct morphemes that must be listed separately. Under this approach, the fact that English also shows similar allomorphy for the third-person singular present tense form of verbs (*eat*[s], *miss*[əz]) and for noun plurals (*cat*[s], *bush*[əz]) would be an even greater adducement to posit a general phonological solution, rather than resorting to unanalyzed allomorphy. Relatedly, Albright and Hayes (2002) propose a phonotactically motivated procedure for collapsing across allomorphs. Under this approach, the learner uses prior knowledge of what kinds of sequences are phonotactically licit in the language to determine whether the observed contextual restrictions on allomorphs (such as $\varnothing \rightarrow$ d/[+voi]_) may be attributed to the desire to avoid phonotactically illegal sequences (**pass*[d], after voiceless). In such cases, the learner posits a devoicing process and attempts to reanalyze all occurrences of the [t] allomorph as instances of [d]. Although both of these approaches have limitations in their implementations, they represent complementary considerations that may drive learners to abstract away from surface differences to arrive at a unified analysis, even in models that do not generally require a single underlying representation for a given morphological marker.

4.4 The Subset Problem in Alternations: Optional Rules

As with phonotactics, learning alternations from positive evidence alone may pose a subset challenge. A particularly interesting case of this, pointed out by Dell (1981), involves the problem of learning whether a rule is optional or obligatory. Dell observes that final obstruent + liquid clusters may optionally be simplified in French: /bukl/ 'buckle' optionally pronounced [buk]. This simplification is not possible for words ending in obstruent + nasal or obstruent + clusters: /ritm/ → [ritm], *[rit] 'rhythm', /fiks/ → [fiks], *[fik] 'fixed'. The challenge for a learner restricted to positive evidence is to determine, based on positive examples like [ritm], that [rit] would not be a grammatical variant. This is fully parallel to the example discussed above in which the learner, presented solely with positive examples of [pa], [ba], and [ma] must infer that [m̥a] is not grammatical. Dell proposes that learners employ an explicit heuristic of adopting the most restrictive grammar possible. As discussed in Section 3.3, one way to implement a restrictiveness bias in Optimality Theory is to favor rankings of Markedness constraints over Faithfulness constraints. It is important to note that in this case, however, the challenge is to demand greater faithfulness in the absence of explicit evidence of alternations (i.e. the learner must assume that nasals and obstruents may not be deleted, but must be pronounced faithfully). It appears that the most general solution to the subset Problem is one that employs a principle such as Entropy (Riggle 2006b) or Maximum Likelihood Estimation (Jarosz), which rely on metrics that bear an invariant relation to restrictiveness, rather than an approach that attempts to regulate Markedness and Faithfulness rankings directly.

4.5 The Gradience Problem in Alternations

As with static phonotactics, alternations do not apply without exception. In many cases, the alternation is lexically restricted: some morphemes consistently undergo them, while others are consistently immune. The Turkish example discussed above could be seen as a case of this: final devoicing and intervocalic voicing are enforced for morphemes such as [kanat] ~ [kanad-ɯ] 'wing-NOM/ACC.', but not for morphemes like /sanat/ 'art' or /etyd/ 'etude'. Numerous studies have used wug tests to explore speakers' knowledge of lexically gradient alternations. In general, it appears that when a process has exceptions and applies with different probability to words of different phonological shapes, speakers' behavior on wug words tracks these differences (Zuraw 2000; Albright, Andrade, and Hayes 2001; Pierrehumbert 2006b; Hayes and Londe 2006; and many others).

As noted above, lexically gradient processes pose a learning challenge because they create inconsistencies that are difficult to capture with a single constraint ranking. Under a theory that attempts to augment underlying representations to reconcile all morphemes with a single grammar, a standard solution is to use diacritics to mark certain morphemes as exceptions to particular rules/constraints. Such a theory attributes no particular significance to the fact that a particular rule has exceptions in some morphemes; the existence of exceptions is simply a static

fact about the lexicon, and no explicit mechanism is provided for speakers to extend gradience to novel items in a wug test. One plausible assumption would be that when speakers are given incomplete information about a novel morpheme in the context of a wug test, they examine the lexicon to assess, the probability of different underlying representations (Schütze 2005); a procedure along these lines is proposed by Harrison and Kaun (2000).

An alternative approach that recognizes and reifies the side-by-side existence of different patterns is to abandon the goal of finding a single consistent grammar, instead allowing morphemes to be associated with different constraint rankings (Ito and Mester 2001b; Anttila 2002a; Inkelas and Zoll 2007; Pater 2000; Becker 2009). For instance, Pater (2009) proposes that when learners are confronted with inconsistent pairs such as *kana**t*** ~ *kana**d**-ɯ* vs. *ety**d*** ~ *ety**d**-y*, they seek to resolve the conflict by finding a constraint that may be ranked differently for different morphemes. In this case, Faithfulness for voicing could be ranked high for words like /etyd/, and ranked low for words like /kanad/. Such proposals make use of the fact that the search space of rankings, while large, is easier to define and search than the space of possible lexicons employing underspecified and augmented underlying representations. In addition, analyses in terms of competing rankings provide a mechanism for encoding the fact that different words behave differently directly in the grammar, providing a natural mechanism for gradient generalization to novel items (Pater 2009). For instance, Becker (2009) proposes that when learners discover that a constraint is variably ranked, they keep track of the number of morphemes that obey each ranking and can use this knowledge to estimate the probability with which a novel morpheme should obey a particular ranking.

Yet another approach to lexically gradient alternations is to use the grammar to encode knowledge of the probability of participating in the alternation, and the lexicon to encode the behavior of individual lexical items. Zuraw (2000) uses the Gradual Learning Algorithm to allow conflicting data from lexically gradient processes to lead to non-categorical rankings, which may then be generalized to novel items at ratios matching the rate of alternation in the training data. In order to capture the fact that existing (known) morphemes are generally consistent in their behavior, it is proposed that speakers rely on memorized word-specific knowledge which blocks the variability that the grammar would otherwise produce. A similar approach can be seen in the minimal generalization model of Albright and Hayes (2003), which encodes competing lexically gradient patterns by means of probabilistic rules, and relies on word-specific knowledge to ensure that known words are inflected consistently.

5 What Doesn't Have to be Learned? The Issue of UG

An important recent development in the study of gradient processes is the possibility that not all statistical trends are equally learnable. For example, Becker,

Ketrez, and Nevins (2011) show that the probability of voicing alternations in Turkish is correlated with a number of features in the surrounding context, such as the place of articulation, the length of the word, and the preceding vowel quality. However, they argue that in this case, wug test data does not mirror the lexical trends as closely as in the examples cited above: Turkish speakers are sensitive to the role of consonant place and word length, but do not appear to take vowel features into account when deciding on the probability of voicing. This highlights the fact that progress in modeling gradient processes is likely to require not only better statistical models of learning from the lexicon, but also a better understanding of which trends speakers choose to encode, and at what level of granularity.

Some of the learning models mentioned above are sharply inductivist, attempting to find the right phonological grammar using very little a priori knowledge, perhaps limited to just a feature inventory and the learning principles themselves. We think the development of such systems is a good research strategy – not because the ultimate right answer to the problem of phonological learning is necessarily a purely inductivist one, but because inductivist approaches can be used to gain insight into UG proposals.

A pioneering contribution in this area is Gildea and Jurafsky (1996), which sought to develop a formal system that, given input/output pairs, could learn appropriate phonological rules to relate the two. They adopted as their baseline algorithm "OSTIA", a procedure for discovering finite-state transducers invented by Oncina, García, and Vidal (1993). Applying OSTIA to English phonological data, Gildea and Jurafsky found that the algorithm could learn versions of rules like flapping only after they had augmented it with three further principles, which are at least tacitly present in almost any phonological theory: "*Faithfulness* (underlying segments tend to be realized similarly on the surface), *Community* (Similar segments behave similarly), and *Context* (Phonological rules need access to variables in their context)" (p. 497). One potential interpretation of this is that the three abstract principles must *necessarily* be part of phonological theory, since learning would be impossible without them. Of course, Gildea and Jurafsky are cautious on this point, since it is possible that some other primitive inductive system might solve the problem as well, or that the three principles might themselves be learnable.

A similar research strategy is adopted in the phonotactic learner of Hayes and Wilson (2008), mentioned above. Hayes and Wilson find that their basic inductive system is defeated by non-local phonological phenomena such as stress and harmony, which can involve segments that are at some distance from one another. They find that such systems can be learned when the phonological theory assumed is augmented to include standard generative phonological formalisms for such phenomena, specifically metrical grids and autosegmental tiers. The crucial difference these representations make is that they provide local formal characterizations of surface-non-local configurations, permitting phonotactic patterns to be learned that would otherwise be inaccessible, being expressible only as hypotheses that occupy huge, unsearchable hypothesis spaces.[17] It is plausible to imagine that a number of other elements of phonological theory would likewise facilitate learning

– and also possible that some proposals actually hinder it, by expanding the hypothesis space with no compensating gain in access to the useful hypotheses.

Learnability studies complement experimental work that seeks to find direct evidence for UG principles. One such type of experiment assesses whether speakers have phonotactic preferences that distinguish forms that are equally unattested in their language: for instance, the form [lbɪf] (monosyllabic) is illegal in English, but has a more severe violation of sonority sequencing principles (e.g. Sievers 1901) than [bdɪf]. Berent *et al.* (2007, 2008; see also Pertz and Bever 1975) find that in various tasks, English speakers act in ways indicating that [lbɪf] is less well-formed [bdɪf], and cautiously suggest that this reflects sonority sequencing as an a priori principle that influences phonetic judgments, independently of whatever phonological principles are learned from exposure to data.

This result can be evaluated further if we use computational learning models. The idea is that perhaps, contrary to initial assumptions, the [lb] – [bd] *is* learnable, being implicit in the ample overt evidence that English onset clusters do respect sonority sequencing in some general sense. Albright (2007b) conducted further experiments on initial sonority sequencing, modeling his results with both an analogical model similar to that of Bailey and Hahn (2001) and his own phonotactic learning algorithm; neither model predicts all the sonority-based acceptability differences in the experimental data, thus tentatively supporting the conclusion that sonority sequencing may embody a priori knowledge. Although this specific conclusion may be overturned by subsequent advances in automated learning, it illustrates a more general principle: computationally implemented learning models provide a concrete estimate of what we can responsibly assume that learners may extract from the data, and guide the researcher towards aspects of phonological patterning that appear to be difficult to extract from the data.

Another important strategy for obtaining grammaticality intuitions that could not have come from the acquisition data is to construct pairs of entirely new miniature languages that differ in crucial respects, and compare people's ability to learn them. The contrasting properties of the language pairs must be uncued (i.e. statistically neutral) in the native language of the experimental subjects. Wilson (2006) set up such an experiment, based on the well-known typological observation that palatalization of velars is favored in the environment before high front vowels relative to before lower ones (Chen 1972). In his experiment, subjects showed some tendency to generalize the rule k → tʃ / ___ e also to cover k → tʃ / ___ i, but not in the opposite direction, a pattern for which their prior (monolingual) experience with English provides no direct evidence. This can be taken, at a very simple level, as a "UG in action" result, but Wilson pursues the issue more intensively by asking what sort of UG, and what kind of learning model might project the result from deeper principles. Wilson's view is that the (perhaps innate) principle at stake is Paradigm Uniformity, taken at the phonetic level (Steriade 2000): speakers are a priori more willing to tolerate alternation between phonetically similar pairs than phonetically distant ones. In the present case, [ke] is further from [tʃe] than [ki] is from [tʃi] (due to the greater burst noise in [ki] than [ke]; Guion 1998), so speakers are a priori more willing to tolerate [ki] ~ [tʃi]

alternation than [ke] ~ [tʃe] alternation. The most striking aspect of Wilson's study is the final step, which is to construct an implemented model of what the experimental subjects were doing when they learned the constructed languages. The model is partly inductive, and partly the application of the innate principle of phonetic Paradigm Uniformity. Formally, Wilson implements this as a constraint-based Maximum Entropy model, in which the weighting of the constraints depends on a prior term that governs the degree to which the constraint weight responds to data during learning. The weights express the final learned grammar, whereas the prior terms express the effect of UG principles on how grammars are learned.

Wilson's work formalizes the idea that UG principles may not always be absolute, but rather can express *learning biases*, whereby the principles guide but do not absolutely dictate the form of the grammar that is shaped on exposure to data. While the bias Wilson examines is phonetic, the variety of cases to be explored and modeled is far wider. Thus, Moreton (2008) offers experimental data suggesting that speakers pick up more easily phonological patterns expressible in terms of identity (here, of vowel quality, or perhaps of height) than other patterns of equal phonetic naturalness (in Moreton's study, a vowel height/consonant voicing correlation).

6 Conclusion

In this chapter, we have attempted to highlight some of the challenges that learners face in analyzing phonological distributions and alternations. In many cases, our current state of knowledge is clearly still quite preliminary, based on schematic and idealized examples. Nonetheless, we believe that significant progress that has been made in formalizing the problem and providing concrete frameworks for solving it since 1955, when Hockett declared that "[w]e know of no set of procedures by which a Martian, or a machine, could analyze a phonologic system" (1955: 147). Furthermore, we anticipate that as computational resources and power expands, current proposals may be subjected to broader and more realistic testing and use of implemented learners will become more widespread, allowing considerations of learnability to play a more central role in guiding phonological theory.

NOTES

1 A side note: for reasons of space we will have nothing to say about a topic of great importance and relevance; that is, *phonetic* learning. By this we include the induction learning of phonological categories (features, segments) from waveforms (studied by, e.g. Mielke 2005b, 2008; Lin 2005a; Maye, Werker, and Gerken 2002), language-specific patterns of phonetic realization (e.g. Keating 1985; Kingston and Diehl 1994), and the vast amount of free variation seen at the phonetic level (as in, for example, coarticulation:

Fowler 1981; Manuel and Krakow 1984; Smith 1992). All phenomena covered here are characterizable at the level of contrasting surface entities.

2 See, for instance, the classic study of Smith (1973), and for a careful overview of more recent work, Demuth's chapter in this volume. For a minority view, claiming a mere physiological basis for child mappings, see Hale and Reiss (1998). An issue that we do not address here is whether children's productions are most appropriately modeled as a distinct grammar (Kiparsky and Menn 1977), or with the same grammar that is used for comprehension (Smolensky 1996b; Pater 2004).

3 On the other hand, the *delearning* of the child's output system does seem to be data-sensitive. Boersma and Levelt (2000) and Curtin and Zuraw (2002) suggest that when the child alters her system to produce outputs closer to adult speech, the process is guided by a preference first to master those marked structures that are more common in the ambient language. For discussion of recent investigations into the relation between frequency and order of productive mastery, see Demuth (this volume, Section 3.6).

4 A terminological note: here we use *input* to refer to a representation that is fed into the grammar to derive a surface representation (= an *output*). We designate input forms with slashes (/ba/), and candidate output forms with square brackets ([ba], [pa]). Our use of these terms is common in theoretical phonology, but must be distinguished from *input* as empirically observed learning data (the input to a model, the input to the child).

5 This approach is similar in spirit to a proposal by McCarthy (2007b) to consider only those candidates that are potentially more harmonic than the fully faithful candidate, an output candidate that faithfully retains all specifications of the input.

6 This can be done by imposing some large, arbitrary upper length on the members of the candidate set.

7 Data gathered as part of the research reported in Albright and Hayes (2003). Average subject ratings on a scale from 1 (worst) to 7 (best) for these five forms were: [kɪp] 5.84, [bzɑɹʃk] 1.50, [fɹɪlg] 2.68, [snɔɪks] 3.00, [pwɪp] 2.89.

8 To be sure, there are sophisticated methods in the *n*-gram literature (see, e.g. Jurafsky, Daniel and Martin 2006: Chapter 6) for interpolating missing *n*-grams. But this is not the same as generalizing over natural classes, an ability that seems crucial in modeling phonotactic knowledge (see Hayes and Wilson 2008: 401; Albright, 2009).

9 For further discussion, see Hayes and Wilson (2008: 402–412). For approaches to nonlocal phonological dependencies not dependent on a prior phonological structure, see Albright and Hayes (2006), Goldsmith and Xanthos (2009).

10 Unlike RCD, the GLA employs a symmetrical reranking scheme, in which loser-preferring constraints are demoted and winner-preferring constraints are promoted. Boersma (1997) argues that this symmetrical strategy is necessary to achieve stasis when variation or errors create conflicting data.

11 For a demonstration of why the GLA can fail, see Magri (2010). Magri's work partly rehabilitates the GLA by offering a revised version, which he proves to converge for non-gradient learning. No proof of adequacy for gradient applications has yet been found, however.

12 For some algorithmic approaches to segmentation into words and morphemes, see Harris (1955), de Marcken (1996), Brent and Cartwright (1996), Goldsmith (2001, 2011), Baroni (2010), and Goldwater (2007).

13 A different approach to the learnability problem posed by hidden structure is to seek observable phonetic differences that would reveal, for example, syllabification of medial

consonant clusters by duration cues (Maddieson 1985; Boucher 1988; Tuller and Kelso 1991) or weight-bearing properties of rhyme consonants (Gordon 2004). Furthermore, hidden structure is only a problem insofar as it actually influences phonological patterning, and in many instances the necessity of hidden structure for this purpose has been questioned or denied; see Prince (1983) and Gordon (2002) on foot structure and stress placement; Steriade (1999a) on syllable boundaries and laryngeal contrast.

14 The proposals in Harrison and Kaun (2000) and McCarthy (2005) are both intended to allow the learner to consider the possibility of extending alternations to morphemes that are not currently known to alternate. We suggest here that a similar strategy would be useful in handling morphemes that are known to alternate but have not yet been analyzed successfully.

15 Support for this idea comes from the fact that Dutch vowel alternations are a relic of a formerly productive pattern of open syllable lengthening (Booij 1995: 88), which has become unproductive/lexically restricted in modern Dutch. The fact that many lexical items have maintained voicing alternations while losing vowel length alternations suggests that the underlying voicing value of individual morphemes was learned successfully, separately from the analysis of vowel length alternations.

16 Blevins (2004) discusses several contexts in which long high vowels do occur in Yokuts and other Yawelmani dialects. The fact that long high vowels are possible in at least some contexts means that the relation between the harmony pattern and the surface phonotactics of Yawelmani is not as direct as it is sometimes portrayed in the literature, and calls into question (but does not preclude) the vowel-lowering analysis reviewed here.

17 Plainly, this is a tentative result, since it depends on the claim that no other learning mechanism would be able to find the crucial generalizations without the a priori provision of tiers and grids; see Goldsmith and Xanthos (2009).

21 Sign Language Phonology

DIANE BRENTARI

1 Introduction

In the last edition of this handbook, Goldsmith (1995) wrote in his introduction,

> The study of signed languages, such as American Sign Language, promises to have a profound effect on phonological theory, and perhaps ultimately on our understanding of what a human language is. The possibilities that emerge from a linguistic system not constrained by the resources of the vocal tract exploit capacities that had until recently been hidden from linguists' view, and the broadened vista that we have today may in retrospect be as significant for the development of linguistics as was the impact of the Western tradition of the study of non-Indo-European languages.

It is now over 50 years since the first work on sign language phonology appeared (Stokoe 1960). The body of work since 1960 has had three basic aims, which will be referred to by the umbrella terms *structure, modality*, and *iconicity*. Under the term *structure* is included all the work that showed that sign languages were natural languages with demonstrable structure at all levels of the grammar including, of course, phonology. Much progress has been achieved toward the aim of delineating the structures, distribution, and operations in sign language phonology, even though this work is by no means over and debates about the segment, feature hierarchies, contrast, and phonological operations continue. For now, it will suffice to say that it is well-established crosslinguistically that sign languages have structures and hierarchical structures analogous to those of spoken languages. Taken together, the five sign language parameters of handshape, place of articulation (where the sign is made), movement (how the articulators move), orientation (the hands' relation towards the place of articulation), and

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

nonmanual behaviors (what the body and face are doing) function similarly to the cavities, articulators, and features of spoken languages.[1] Despite their different content, these groups of features in sign languages are subject to operations that are similar to those of their spoken language counterparts. These broad-based similarities must be seen, however, in the light of important differences due to modality and iconicity effects on the system. *Modality* addresses the effect of peripheral systems (i.e. visual/gestural vs. auditory/vocal) on the very nature of the phonological system that is generated. *Iconicity*, which is a specific type of modality effect, refers to the non-arbitrary relationships between form and meaning (Brennan 1990, 2005), particularly visual/spatial iconicity in the case of sign languages.

This chapter will be structured around the three themes of structure, modality, and iconicity because these issues have been studied in sign language phonology (indeed, in sign language linguistics) from the very beginning. Section 2 will outline the phonological structure of signed languages, adding new material to what was covered in Brentari (1995), in the Handbook's first edition. In Section 3, several typological facts about signed languages regarding word-level phonotactics will be used as an example to illustrate communication modality effects. What is it about the visual nature of signed language that creates this typological niche? This can be answered in terms of signal processing, and experimental evidence from word segmentation supports a modality explanation for the typological facts. Section 4 will focus on iconicity. It will be argued that this concept is not in opposition to arbitrariness, but that iconicity co-exists along with other factors – such as ease of perception and ease of production – that contribute to the inventory of sign language phonological elements. Iconicity is pervasive in the inventory of phonological elements in signed languages, but the distribution of these elements is arbitrary, and it is in the distribution of elements that the true test of phonology lies.

2 Structure

2.1 General Organization of Sublexical Structure

The structure in Figure 21.1 shows the three basic manual parameters – Handshape (HS), Place of Articulation (POA), and Movement (M) – in a hierarchical structure, that of the Prosodic Model (Brentari 1998). This structure presents a fundamental difference between signed and spoken languages. Besides the different featural content, the most striking difference between signed and spoken languages is the hierarchical structure itself – i.e. the root node at the top of the structure is an entire lexeme, a stem, not a consonant- or vowel-like unit.

Both signed and spoken languages have impressive amounts of simultaneous structure, but the representation in Figure 21.1 encodes the fact that a high number of features are specified only once per lexeme in sign languages. This idea will be described in detail below. Since the beginning of the field there has been debate about how much to allow the simultaneous aspects of sub-lexical sign structure to dominate the representation: whether sign languages have the

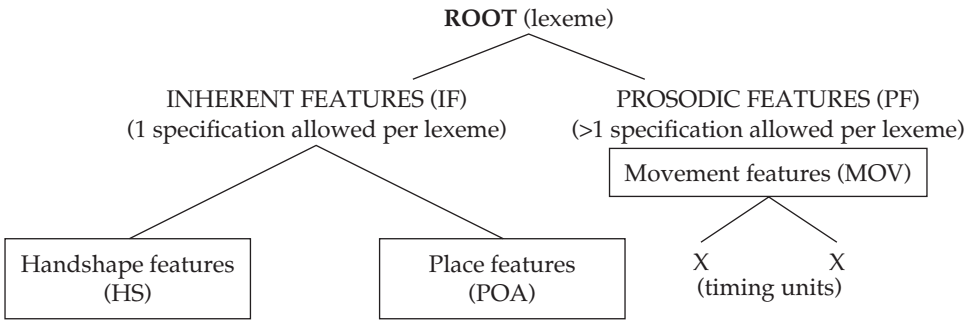
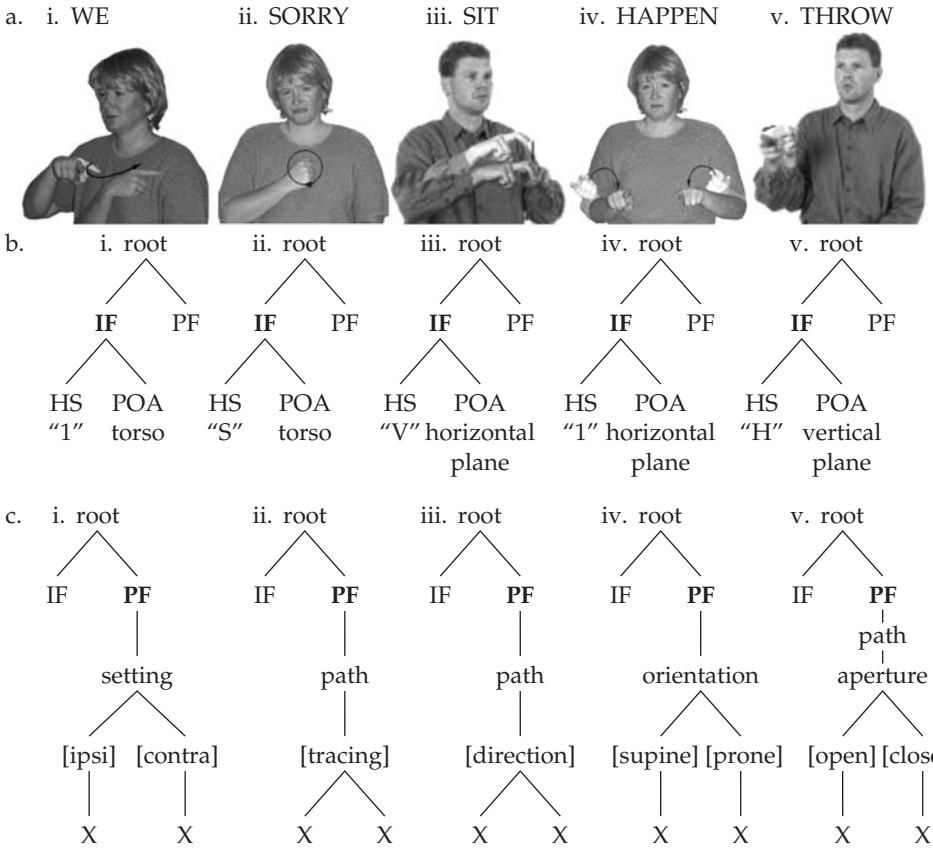


Figure 21.1 The hierarchical organization of handshape, movement, and place of articulation in a sign, according to the Prosodic Model (Brentari 1998).

same structures and structural relationships as spoken languages, but with lots of exceptional behavior, or a different structure entirely. A proposal such as the one in Figure 21.1 is for a different structure, a bold move not to be taken lightly. Based on a wide range of available evidence, however, it appears that the simultaneous structure of signs is indeed more prevalent in signed than in spoken languages; see Sandler (1989), Brentari (1990a, 1998), Channon (2002), van der Hulst (2000), Sandler and Lillo-Martin (2006).[2]

The general concept of "root-as-lexeme" present in most phonological models of sign language phonology accurately reflects the fact that sign languages typically specify many features just once per lexeme. These are the *inherent features* of the Prosodic Model presented in Figure 21.1. Looking at the signs illustrated in Figure 21.2a and their related Place of Articulation and Handshape features in Figure 21.2b, one sees this point clearly. (The major Place of Articulation and Handshape have many feature specifications, all of which remain the same, but for our purposes, it is enough to show them all as a composite symbol in single quotes – e.g. '1', 'S', etc.). Regarding Place of Articulation, even though it looks like the hand starts and stops in a different places in each sign, the major region where the sign is articulated is the same – the *torso* in WE and SORRY, the *horizontal plane* in front of the signer in SIT and HAPPEN and the *vertical plane* in front of the signer in THROW. These are examples of contrastive places of articulation within the system. With regard to Handshape, there is just one Handshape in the first three signs – WE, SORRY, and SIT. The Handshape does not change at all throughout articulation of the sign. In the last sign the two fingers change from closed to open, but the *selected fingers* used in the handshape do not change. The opening is itself a type of movement, which is described below in more detail. For further elaboration on the separation of selected fingers and aperture, see van der Hulst (1995) and Brentari (1998).

In the space provided here full argumentation for the features and their positions in the structure will not be discussed, but the main point is that relatively more features in sign languages are specified just once per lexeme than is the case with the



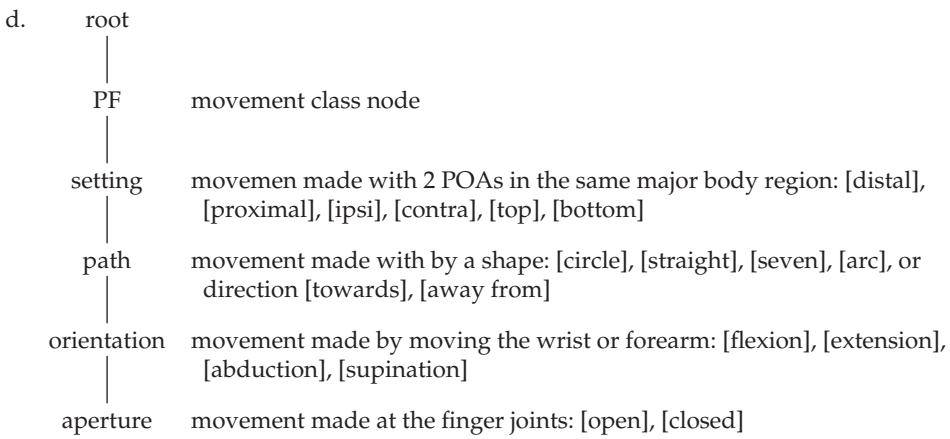


Figure 21.2 Illustrations of signs (a) their elements that are specified one time per lexeme (b), their elements that change during the production of the sign (c), and the structure of prosodic features [movement] (d).

features of spoken languages. Tone in tonal languages, and features that harmonize across a lexeme (e.g. vowel features and nasality) behave this way in spoken languages, but fewer features seem to have this type of domain in spoken than in signed languages. And when features do operate this way in spoken languages, it is not universal for all spoken languages. In sign languages a larger number of features operate this way and they do so relatively across sign languages.

The *prosodic features* in Figure 21.1 are those describing movements within the sign, such as the aperture change just mentioned. These features allow for changes in their values within a single root node (lexeme) while the inherent features do not, and this phonological behavior is the justification for isolating the movement features on a separate autosegmental tier. The trees in Figure 21.2c demonstrate different types of movement features for the signs in Figure 21.2a, and the whole movement feature structure is shown in more detail as the prosodic features (PF) node in Figure 21.2d.

Each specification indicates a different joint (or joints) that can be responsible for a movement – shoulder, elbow, wrist, and hand – progressing top to bottom in the representation from the more *proximal* joints of the arm to the more *distal* joints of the hand. In other words, the shoulder articulating the *setting* movement in WE is located closer to the center of the body than the elbow that articulates a *path* movement in SORRY and SIT. A sign having an *orientation change* (e.g. HAPPEN) is articulated by the wrist, a joint that is even further away from the body's center, and an *aperture change* (e.g. THROW), is articulated by the hand, furthest away from the center of the body. Notice that it is possible to have two simultaneous types of movement articulated together; the sign THROW has a path movement and an aperture change.

The timing slots projected from the prosodic structure are also shown in Figure 21.2c. The inherent features do not generate timing slots at all, only movement features can do this. When two movement components are articulated simultaneously as in THROW, they align with one another and only two timing slots are projected onto the timing tier. The movement features have been described in detail here because they play an important role in the sign language syllable, discussed in the next section.

2.2 The Syllable[3]

The syllable is as fundamental a unit in signed as it is in spoken languages. One point of nearly complete consensus across models of sign language phonology is that the movements are the nuclei of the syllable. This idea has its origin in the correlation between the function of movements and the function of vowels in spoken languages (Liddell 1984), but this was developed into a theory of syllable structure by Brentari (1990a) and Perlmutter (1992). The arguments for the syllable are based on its importance to the system. They are:

2.2.1 The Babbling Argument Petitto and Marentette (1991) have observed that a sequential dynamic unit formed around a phonological movement appears in

young Deaf children at the same time as hearing children start to produce syllabic babbling. Because the distributional and phonological properties of such units are analogous to the properties usually associated with syllabic babbling, this activity has been referred to as "manual babbling." Like syllabic babbling, manual babbling includes a lot of repetition of the same movement, and also like syllabic babbling, manual babbling makes use of only some of the phonemic units available in a given sign language. The period of manual babbling develops without interruption into the first signs (just as syllabic babbling continues without interruption into the first words in spoken languages). Moreover, manual babbling can be distinguished from excitatory motor hand activity and other communicative gestures by its rhythmic timing, velocity, and spectral frequencies (Petitto 2000).

2.2.2 The Minimal Word Argument This argument is based on the generalization that all well-formed (prosodic) words must contain at least one syllable. In spoken languages a vowel is inserted to insure well-formedness, and in the case of signed languages a movement is inserted for the same reason. Brentari (1990b) observed that ASL signs without a movement in their input, such as the numeral signs 1 to 9 and fingerspelled letters containing no movement in their base form, add an epenthetic movement when used as independent word. Jantunen (2007) observed that the same is true in Finnish Sign Language (FinSL), and Geraci (2009) has observed a similar phenomenon in Italian Sign Language (LIS).

2.2.3 Evidence of a Sonority Hierarchy Many researchers have proposed sonority hierarchies based 'movement visibility' (Corina 1990; Perlmutter 1992; Sandler 1993; and Brentari 1993). Such a sonority hierarchy is built into the prosodic features' structure in Figure 21.2d since movements represented by the more proximal joints higher in the structure are more visible than those articulated by the distal joints represented lower in the structure. For example, movements executed by the wrist are typically larger and more easily seen from further away than those articulated by the hand.

In a study of fingerspelled words used in a series of published ASL lectures on linguistics (Valli and Lucas 1992), Brentari (1994) found that fingerspelled forms containing strings of eight or more handshapes representing the English letters were reduced in a systematic way to forms that contain fewer handshapes. The remaining handshapes are organized around just two movements. This is a type of nativization process by which such forms conform to sign language word-level phonotactics by having no more than two syllables. Crucially, the movements retained were the most visible ones, argued to the be most sonorous ones – e.g. movements made by the wrist were retained while aperture changes produced by the hand were deleted. Figure 21.3 contains an example of this process: the carefully fingerspelled form P-H-O-N-O-L-O-G-Y is reduced to the letters underlined, which are the letters responsible for the two wrist movements.

Some researchers have considered this a manifestation simply of visual 'loudness' (Crasborn 2001; Sander, and Lillo-Martin 2006). While it is true both in spoken and signed languages that more sonorous elements of the phonology are louder

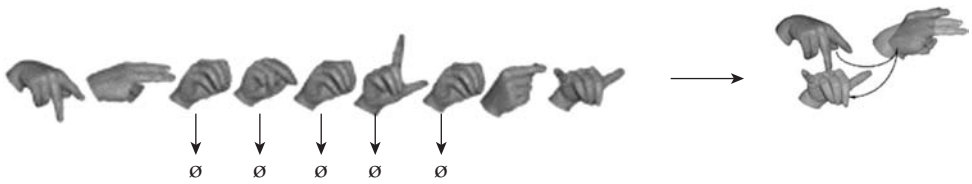


Figure 21.3 Evidence of the Sonority Hierarchy. Nativization of finger-spelled word P-H-O-N-O-L-O-G-Y.

than less sonorous ones (/a/ is louder than /i/; /l/ is louder than /b/, etc.), the evidence from the nativization of fingerspelled words indicates that sonority has infiltrated the word-level phonotactics of sign languages.

2.2.4 Evidence for Light vs. Heavy Syllables Further evidence for the syllable comes from a division between those movements that contain just one movement element (features on only one tier of Figure 21.2d are specified), which behave as *light* syllables (e.g. WE, SORRY, and SIT in Figure 21.2 are light) vs. those that contain more than one simultaneous movement element, which behave as heavy syllables (e.g. THROW in Figure 21.2). It has been observed in both ASL and FinSL that a process of nominalization by movement reduplication can occur only to forms that consist of a light syllable (Brentari 1998; Jantunen 2007). In other words, holding other factors constant, there are signs, such as SIT and THROW, that have two possible forms: a verbal form with the whole sequential movement articulated once, and a nominal form with the whole movement articulated twice (in a restrained manner). The curious fact is that the verb SIT has such a corresponding nominal form (CHAIR), while THROW does not. These facts can be explained by the generalization that the set of forms that allow reduplication have just one simultaneous movement component, and are *light* syllables, while those that disallow reduplication, such as THROW, have two or more simultaneous movement elements and are therefore *heavy*.

These sections on word-level phonology and the syllable show clearly that sign languages have all of the elements one might expect to see in a spoken language phonological system, yet their organization and content is somewhat different. What motivates this difference? One might hypothesize that this is in part due to the visual/gestural nature of sign languages, and this topic of modality effects will be taken up in Section 3 of this chapter.

2.3 Prosodic Structure

One area on which sign language phonology has made remarkable progress in the last 15 years has been prosodic structure. Miller (1996), Nespor and Sandler (1999), Sandler (1999), Wilbur (1999), Brentari and Crossley (2002), and Sandler and Lillo-Martin (2006) have worked on various prosodic units, including the

prosodic word (P-word), phonological phrase (P-phrase), and intonational phrase (I-phrase). As in spoken languages, phonological constituents are not isomorphic with their morphological or morpho-syntactic counterparts, but they are related to them in important ways demonstrated for spoken languages in work such as Nespor and Vogel (1986), Selkirk (1984b), and Truckenbrodt (1999). Cues of prosodic structure in signed languages are listed in (1). Prosodic cues are neither all domain cues nor all boundary cues, but domain cues are more prevalent, and more prevalent than domain cues in spoken languages. The P-word cues are all domain cues; in other words, the outputs of Nonmanual Spreading and Handshape Assimilation, discussed below, result in a single value for these properties across the entire P-word. Some of the cues at the level of phrase are domain cues (Nondominant Handshape Spread) and some are boundary markers (Lengthening). I-Phrase cues are mixed as well; blinks are boundary markers, but many of the nonmanual cues that carry syntactic properties, such as a brow raise over an entire 'yes/no' question, are domain cues. Sandler and Lillo-Martin (2006) argued that regardless of the sources of the nonmanual behavior (syntactic, adverbial, etc.,) their domains and distribution are prosodic in nature. Nespor and Sandler (1999) have distinguished between the properties of movement, such as those discussed for the P-Phrase, by calling them "rhythm cues" and calling the cues on the face "intonation," but both rhythm and intonation cues are undoubtedly prosodic. Suffice it to say that many of the important prosodic cues have been noted for several signed languages, but we have yet to understand fully the extent and distribution of all of their uses crosslinguistically.

(1) Prosodic cues in sign languages

- a. Prosodic word:
  - i. Assimilation of the handshape occurs across a cliticized pronoun (Liddell and Johnson 1989; Sandler 1999).
  - ii. Spreading of mouthing (borrowed lip patterns of borrowed spoken words) occurs across two morphological words in a prosodic word (Boyes Braem 1999, 2001; Brentari and Crossley 2002).
  - iii. Coalescence of dominant handshape across two morphological words creating one phonological word (Sandler 1999).
- b. Phonological phrase:
  - i. Spreading of the nondominant handshape (Nespor and Sandler 1999).
  - ii. Phrase-final lengthening (Perlmutter 1992; Miller 1996; Brentari 1998).
- c. Intonational phrase:
  - i. Eyeblinks at the right edge of I-Phrases (Wilbur 1994; Nespor and Sandler 1999).
  - ii. Changes in leans from left to right in the signing space (Wilbur and Patschke 1998; Boyes Braem 1999).[4]
  - iii. Resetting of all nonmanual behaviors (Nespor and Sandler 1999).

In Figure 21.4 we see examples of markers of each type of constituent. Figure 21.4a, i shows prosodic word markers, using an example from ISL for Handshape

I READ

a, i. P-word: handshape assimilation (ISL)

COLD SHOULDER

a, ii. P-word: Non-manual spread (ASL)

FROM C-O-D-A

b. P-phrase: Non-dominant hand spread (ASL)

ENGLISH ASL DIFFERENT

LANGUAGES (1^{st} position) (2^{nd} position) $[\text{blink}]]_{IP}$

WOW MIND BLOWOUT

c. Blinking at I-Phrase boundaries

Figure 21.4 Prosodic constituent markers in sign languages: (a) Prosodic Word markers in ISL and ASL; (b) Phonological Phrase markers in ASL; (c) Intonational Phrase markers in ASL.

Assimilation – described in (1a, i) of the list above – which appears in forms containing clitics, as well as compounds (Liddell and Johnson 1989; Sandler 1999, Sandler and Lillo-Martin 2006). Notice that the "index finger" handshape, which is used for the first person pronoun when it appears as an independent word, does not appear on the right hand; instead this cliticized form exhibits the same handshape as the following sign READ (the two-finger, "V" handshape); this indicates that the two forms are one P-word. In Figure 21.4a, ii – described in (1a, ii) above – we see an example of Nonmanual Spread (Boyes Braem 1999; Brentari and Crossley 2002). In this case, the mouth posture is the same across the two morphological words COLD and SHOULDER that have become one prosodic word in the expression "cold-shoulder" with the same meaning as the English expression.

Phonological phrase markers are less reliable than P-word or I-Phrase markers, but there are two worth mentioning. In Figure 21.4b is an example of a Handshape Assimilation of the Nondominant hand,[5] described in (1b, i) above. Notice that the left hand remains in the signing space while signing FROM C-O-D-A on the right hand (the 'c-o-d-a' is fingerspelled). Notice also that there is a change in the posture of the mouth between the signs FROM and C-O-D-A, so as discussed above, this cannot be judged to be a single prosodic word; if it were, there would be only one mouth posture. This phrase was extracted from the sentence "From a C-O-D-A]$_{\text{P-phrase}}$ [it is a different story]." Brentari and Crossley (2002) found several different uses of H2 spread; it occurs in compounds and idiomatic expressions (such as COLD-SHOULDER, just mentioned) so it was not necessarily the most reliable indicator of a P-phrase. Phrase-final lengthening is slightly more reliable; that is, comparing a phrase-internal and a phrase-final form of the same sign, there will be at least a 1.5 increase in duration in the phrase-final sign. Several different units have been suggested for lengthening. Perlmutter (1992) applied lengthening to the segment, Liddell (1978), Wilbur and Nolen (1986) and Miller (1996) suggested that movement (i.e. syllables) might be the relevant unit. Tang *et al.* (2010) applied the 1.5 formula to the entire sign.

The three markers of an I-Phrase given in (1c) are all optional; we will discuss only one of these markers – eyeblinks. Wilbur (1994) and Nespor and Sandler (1999) have argued that inhibited periodic eyeblinks are an I-Phrase boundary marker. These are short blinks that occur either just after or slightly overlapping with the final sign in an I-Phrase. The English translation of the passage in Figure 21.4c is: "[English and ASL are different languages]$_{\text{IP}}$ [WOW! It blows my mind.]$_{\text{IP}}$" While fairly consistent as a I-Phrase marker across sign languages, there is evidence of crosslinguistic variation in the use of this prosodic cue. Tang *et al.* (2010) studied blinks in Hong Kong Sign Language (HKSL), Japanese Sign Language (JSL), Swiss German Sign Language (DSGS), and ASL. Across all four signed languages blinks were used to mark the right edges of I-Phrases, just as Wilbur found in ASL, but blinks co-occurred with different cues in JSL than in the other three sign languages. In ASL, DSGS, and HKSL blinks were associated with lengthening, while in JSL, blinks were associated more often with head nods. Even more importantly, it was found that, holding signing rate constant, there

are different blink rates across signed languages, suggesting that even though all four sign languages used blinks to mark I-Phrases, some sign languages may use blinks to mark smaller constituents as well.

Before concluding this section on sign language prosody it is worth mentioning the use of similar prosodic cues in signers and in speakers. There is an undeniable overlap in the content and use of prosodic cues in co-speech gesture and in signed languages, and this topic is receiving attention in the literature. The intentions (precursors to speech acts) of 1-year-old, nonsigning, toddlers at the one-word stage of language acquisition are better understood by caregivers when their gestures and speech cues coincide with one another in timing and content (Balog and Brentari 2008). In terms of timing, there is evidence that pitch accents coincide with manual gestures in adults (Loehr 2004) and that they have an effect on the perceived prominence of accented syllables (Krahmer and Swerts 2007). Furthermore, nonsigning 7.5-month-old-infants are more attentive to material in which the prominence expressed visually on the face is in synchrony with the peaks of prosody of the spoken signal; infants at this age are able to segment words from the speech stream better in such a context (Hollich *et al.* 2005). Finally, several studies have also demonstrated that nonsigners are sensitive to the specific prosodic cues of signed languages. In two studies adult nonsigners were able to perceive the presence of I-Phrase boundaries in a reliable fashion (Fenlon *et al.* 2007; Brentari *et al.* 2011); Brentari *et al.* (2011) also showed a similar result in 9-month old hearing (non-signing) infants.

The use of the prosodic cues mentioned in (1) is clearly arbitrary in sign languages; both distributional evidence and neuro-imaging studies confirm this. Facial expressions that are grammatical are left lateralized in signers but not for non signers (McCullough *et al.* 2005). Nevertheless, the nature of "gestural competence" in nonsigners can also be informed by the work on sign languages, since there is an overlap in the use of visual/gestural cues as prominence markers and boundary markers in signed and spoken languages.

3 Modality Effects (The Effects of Communication Mode)

The modality effects described here refer to the influence that the phonetics (or communication mode) used in a signed or spoken medium have on the very nature of the phonological system that is generated. How is communication modality expressed in the phonological representation? A few differences between sign and speech that affect the phonetics are: (i) a signed word takes longer to articulate than a spoken one; (ii) an auditory signal is treated differently than a visual signal in processing; (iii) signed words can draw on visual similarities with the entities represented much more easily than spoken words can. I am claiming that strong statistical tendencies in signed and spoken languages (not absolutes) emerge because the communication mode contributes significantly to even the most abstract of phonological structures. Word shape will be used as an example

of how modality effects ultimately become reflected in phonological and morphological representations.

3.1 Word Shape

In this section, the differences in the shape of the canonical word in signed and spoken languages will be described, first in terms of typological characteristics alone, and then in terms of factors due to communication modality. "Canonical word shape" refers to the preferred phonological shape of words in a given language. For an example of such canonical word properties, many languages, including the Bantu language Shona (Myers 1987) and the Austronesian language Yidiny (Dixon 1977), require that all words be composed of binary branching feet. With regard to statistical tendencies at the word level, a preferred canonical word shape is also exhibited by the relationship between the number of syllables and morphemes in a word, and it is here that signed languages differ from spoken languages. Signed words tend to be monosyllabic (Coulter 1982); that is, referring back to the movement structure in Figure 21.2d, the Stokoe *et al.* (1965) dictionary shows that approximately 83% of the lexical entries are composed of single sequential movements. And, unlike spoken languages, signed languages have a proliferation of monosyllabic, polymorphemic words because most affixes in sign languages are feature-sized and are layered simultaneously onto the stem rather than concatenated (see also Aronoff *et al.* (2005) for a discussion of this point).

This relationship between syllables and morphemes is a hybrid measurement, which is both phonological and morphological in nature, in part due to the shape of stems and in part due to the type of affixal morphology in a given language. A spoken language, such as Hmong, contains words that tend to be monosyllabic and monomorphemic with just two syllable positions (CV), but 39 consonants and 13 vowels. The consonant inventory includes as secondary articulations voiced and voiceless nasals, pre- and post-nasalized obstruents, and lateralized obstruents, and the inventory of vowels includes monophthongs and, diphthongs, and seven contrastive tones, including simple and contour tones (Golston and Yang 2001; Andruski and Ratliff 2000). A language, such as West Greenlandic, contains stems of a variety of shapes and a rich system of affixal morphology that lengthens words considerably (Fortescue 1984). In English, stems tend to be polysyllabic, and there is relatively little affixal morphology. In sign languages, words tend to be monosyllabic, even when they are polymorphemic. An example of such a form – re-presented from Brentari (1995: 20) – is given in Figure 21.5; this form means "two bent-over upright beings advance-forward carefully side-by-side" and contains at least six morphemes in a single syllable. The agreement forms in Figure 21.7 and all of the classifier constructions in Figure 21.10 (discussed later in Section 4) are monosyllabic as well. There is a large amount of affixal morphology, but most of these affixes are smaller than a segment in size; hence, polysyllabic and monosyllabic words are typically not different in word length. In (2), a chart schematizes the canonical word shape in terms of the number of morphemes and syllables per word.

Figure 21.5 An example of a monosyllabic, polymorphic form in ASL: "two bent-over upright-beings advance-forward carefully side-by-side."

(2) Canonical word shape according to the number of syllables and morphemes per word

	Monosyllabic	*Polysyllabic*
Monomorphemic	Hmong	English, German, Hawaiian
Polymorphemic	sign languages	West Greenlandic, Turkish, Navajo

This typological fact about signed languages has been attributed to communication modality, as a consequence of their visual/gestural nature. This logic predicts that simultaneous systems with a high degree of simultaneous structure both at the level of the stem and in the affixal morphology should be relatively rare in spoken languages and common in sign languages, but notions such as "high degree" and "rare" must be cashed out using available corpora to be viable.

Without a doubt, spoken languages have simultaneous phenomena in phonology and morphophonology such as tone, vowel harmony, nasal harmony, and ablaut marking (e.g. the past preterit in English (*sing*-pres./*sang*-preterit, *ring*-pres./*rang*-preterit), and even person marking in Hua indicated by the [±back] feature on the vowel (Haiman 1979)). There is also nonconcatenative morphology found in Semitic languages, which is another type of simultaneous phenomenon, where lexical roots and grammatical vocalisms alternate with one another in time. Even collectively, however, this does not approach the degree of simultaneity in signed languages, because, as mentioned in Section 2, many features are specified once per stem to begin with – one handshape, one movement, one place of articulation. Add to that the fact that the morphology is feature-sized and layered onto the same monosyllabic stem, adding additional features but no more linear complexity, and the result is that sign languages have two sources of simultaneity – one phonological and another morphological. I would argue that it is this combination of these two types of simultaneity that causes signed languages to occupy this

typological niche. Many researchers since the 1960s have observed a preference for simultaneity of structure in signed languages, but for this particular typological comparison it was important to have understood the nature of the syllable in signed languages; that is, the syllable is based on the sign's movement component, as discussed in Section 2.

In the next section experimental evidence is presented showing that the typological fact about signed language word shape is indeed due to the visual/gestural modality, and not accidental or based on language experience.

3.2 Simultaneous vs. Sequential Processing

Consider the typological fact just described about canonical word shape from the perspective of the peripheral systems involved and their particular strengths in signal processing. "Simultaneous processing" is a cover term for our ability to process various input types presented roughly at the same time (e.g. pattern recognition, paradigmatic processing in phonological terms); "sequential processing" is our ability to process temporally discrete inputs into temporally discrete events (e.g. ordering and sequencing of objects in time, syntagmatic processing in phonological terms). Despite the fact that many aspects of simultaneous and sequential processing take place in both vision and audition, there are differences in the inherent strengths built into the design of the physiological visual and auditory peripheral systems, as outlined in (3).

(3) Differences between signal processing in vision and audition (from Brentari 2002)

	Vision	*Audition*
Speed of signal transmission	299, 300 km/sec	.33 km/sec
Peripheral temporal resolution	25–30 ms	2 ms
Spatial arrangement information	Peripheral	Non-peripheral

In general, the advantage in sequential processing goes to audition, while the advantage in simultaneous processing goes to vision. For example, the time required for a subject to detect temporally discrete stimuli is a sequential processing task. The time required for detection in vision vs., audition is different. In vision this period is called the "threshold of flicker fusion" (Chase and Jenner 1993), and in audition the "threshold of temporal resolution" (Kohlrausch, Püschel, and Alphei 1992). Humans can temporally resolve auditory stimuli when they are separated by an interval of only 2 milliseconds (Green 1971; Kohlrausch, Püschel, and Alphei 1992), while the visual system requires at least 20 milliseconds to resolve visual stimuli presented sequentially (Chase and Jenner 1993). Meier (2002) also discusses the ability to judge duration and rate of stimulus presentation; hence the advantage in temporal processing goes to audition.

Regarding simultaneous processing, one effect of the speed of light transmission on the perception of objects is that vision can take advantage of light waves

reflected not only from the target object, but also by secondary reflection from other objects in the environment onto the target object, thereby making use of visual "echo" waves. These secondary reflections are perceived simultaneously with the waves directly reflected from the target object to the retina, enhancing its three-dimensional quality (Bregman 1990). This same echo phenomenon in audition is available to the listener only much later. The result of this effect is that vision allows a more three-dimensional image to be available more quickly due to properties of the signal itself (light vs. sound waves). Moreover, the localization of visual stimuli is registered at the retina and lens, physiologically the most peripheral component of the visual system, while the spatial arrangement of auditory stimuli is resolved at the cortical level and can only be inferred by temporal and intensity differences of the signal between the signals of the two ears (Bregman 1990).[6] Meier (2002) also discusses the transmission property of bandwidth, which is larger in vision, and spatial acuity, which is the ability to accurately pinpoint an object in space (Welch and Warren 1986). All of these factors award the advantage of spatial resolution to vision.

One might, therefore, expect words in signed and spoken languages to exploit the advantages available to the system.

3.3 Word Segmentation is Grounded in Communication Modality

If the typological difference between words in signed and spoken language described in Section 3.1 is deeply grounded in communication modality it should be evident without language experience. From a psycholinguistic perspective, the phenomenon of word shape can be fruitfully explored using word segmentation tasks, because it can address how language users with different experience handle the same types of items. We discuss two such studies in this section.

The cues that people use to make word segmentation decisions are typically put into conflict with each other in experiments to determine their relative salience to perceivers. Word segmentation judgments in spoken languages are based on (i) the rhythmic properties of metrical feet (syllabic or moraic in nature), (ii) segmental cues, such as the distribution of allophones, and (iii) domain cues, such as the spreading of tone or nasality. Within the word, the first two of these are "linear" or "sequential" in nature, while domain cues are simultaneous in nature and coextensive with the whole word. These cues have been put into conflict in word segmentation experiments, and it has been determined crosslinguistically that rhythmic cues are more salient when put into conflict with domain cues or segmental cues (Vroomen *et al.* 1998; Jusczyk *et al.* 1993a, 1999; Houston *et al.* 2000). Rhythm cues are linear alternations (e.g. '**chil.**dren, '**break.**fast), but they unfold more slowly than other linear changes, such as segmental alternations. Rhythm cues also require less specialized knowledge about the grammar; that is, there are only a few prosodic differences that are logically possible if we assume that there is at least one prominent syllable in every word (two-syllable words have three possibilities; three-syllable words have seven possibilities). Segmental

alternations, on the other hand, such as knowing the allophonic form that appears in coda vs. onset position, requires language-particular knowledge at a rather sophisticated level, though infants master it some time between 9 and 12 months of age (Jusczyk *et al.* 1999).

Word-level phonotactic cues are available for sign languages as well, and these have also been used in word segmentation experiments.[7] This has already been introduced in Section 2, when discussing which features have one value per word (the inherent features) and which ones can change (prosodic features). Some word-level phonotactics are described in (4):

(4) Word-level phonotactics[8]

a. Handshape: within a word selected finger features do not change; aperture features may change.
b. Place of articulation: within a word major place of articulation features may not change; setting features (minor place features) within the same major body region may change.
c. Movement: within a word repetition of movement is possible, or "circle+straight" combinations (*straight+circle).

The research question is: What properties of sign language play more of a role in sign language word segmentation – the ones that do not change (the simultaneous ones) or the ones that do (i.e. those that are sequential in nature)? These cues were put into conflict with one another in a set of balanced nonsense stimuli that were presented to signers and nonsigners. The use of a sequential cue might be, for example, noticing that the open and closed aperture variants of handshapes are related, and thereby judge a form containing such a change to be one sign. The use of a simultaneous strategy might be, for example, to ignore sequential alternations entirely, and to judge every handshape as a new word. The nonsense forms in Figure 21.6 demonstrate this. If a participant relied on a sequential strategy, Figure 21.6a would be judged as one sign because it has an open and closed variant of the same handshape, and Figure 21.6b would be judged as two signs because it contains two distinctively contrastive handshapes (two different selected finger groups). If, on the other hand, a participant relied on a simultaneous strategy, both of the signs in Figure 21.6 would be judged as two signs.

In these studies, four groups of subjects took part in two experiments. In one study, groups of native users of ASL and English participated (Brentari 2006), and in another study there were groups of native users of ASL, Croatian Sign Language (HZJ), English, and Croatian (Brentari *et al.* 2011). All were administered the same word segmentation task in which the participants were asked to judge whether controlled strings of nonsense stimuli based on ASL words were one sign or two signs. It was hypothesized that speakers without signing experience might exploit linear cues of the signal since sequential cues, such as strong/weak alternation in foot structure, are so important in spoken languages. It was also hypothesized that that signers would use their language-particular knowledge to judge Figure 21.6a as one sign and 21.6b as two signs. Hypothesis 1 was not confirmed;

a. "1 sign" b. "2 signs"

Figure 21.6 Examples targeting handshape in the word segmentation experiments. An item that should elicit a "1-sign" judgment based on a segmental strategy (a), and another (b) with an expected "2-sign" judgment. Both (a) and (b) should elicit a "2-sign" judgment if a general "one value = 1 word" strategy is used since there is a handshape change in both cases.

there was no significant difference between the signing and non-signing groups' performance. The strongest result was that *both* speakers and signers used a domain/simultaneous strategy when segmenting the sign stream into words across all major feature types, but especially in the movement parameter. Hypothesis 2 was confirmed only in the Handshape parameter, and only for signers. Otherwise, both signers and nonsigners used a "1 value = 1 word" strategy overall – a domain strategy – despite the specialized grammatical knowledge for movement and location within words (in the ASL signers' case).

The conclusion drawn from these word segmentation experiments is that modality plays a powerful role in word segmentation. Domain (simultaneous) cues are stronger in sign languages than sequential cues. Since sequential cues are stronger in spoken languages, it might be reasonable to expect that this might, at least in part, be due to modality as well.

This does not mean, however, that a strategy that is dispreferred by the relevant modality is never employed. On the contrary, many spoken languages use domain effects of tone or nasality to signal word boundaries, as has been noted previously, and sequential effects within the word in signed languages have also been noted (see also Aronoff *et al.* 2005). It does mean that, when faced with a new type of linguistic string, the modality will play a role in segmenting it. Incorporating this factor into the logic of phonological architecture would help explain why certain structures, such as the trochaic foot, may be so powerful a cue to word learning in infants (Jucszyk *et al.* 1999).

3.4 The Reversal of Segment to Melody

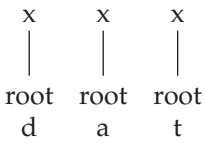
A final modality effect is the organization of melody features to skeletal segments in the hierarchical structure in Figure 21.1. Notice that the content of the

features predicts the skeletal structure in sign languages; as a consequence the features have a higher position and timing slots a lower position in the structure: namely, the reverse of what occurs in spoken languages. Skeletal slots are predictable and hence have a lower position in the structure.[9] The reason timing units are higher in the hierarchical structure of spoken languages is because they can be contrastive. Length is not contrastive in any known sign language. But why would this be the case? One reason has already been mentioned in Section 3.2: audition has the advantage over vision in making temporal judgments, so it makes sense that the temporal elements of speech have a powerful and independent role in phonological structure with respect to the melody (i.e. the featural material). One consequence of this might be that the skeletal tier, containing either segments or moras, is more heavily exploited to produce contrast within the system.

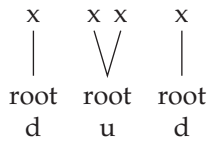
In spoken languages, affricates, geminates, long vowels, and diphthongs demonstrate that the number of timing slots must be represented independently from the melody, even if the default case is one timing slot per root node. Examples of the segment-root-feature structure for English are given in (5). A schema for the root-feature-segment structures for both spoken and signed languages is given in (6).[10,11]

(5) spoken language phonology – root-segment ratios (English)

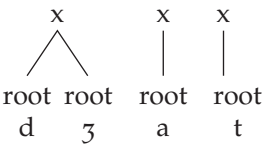
a. 1:1 [a] in 'dot'



b. 2:1 [u:] in 'dude'



c. 1:2 [ʤ] in 'jot'



(6) Organization of phonological material in signed vs. spoken languages

a. Spoken languages

b. Signed languages

To conclude this section on modality, researchers working on signed languages confront such issues as the ones described here constantly, since linguistic theories have been developed largely for spoken languages, and as a result we question the influence of the visual modality at every turn. Data from signed languages push the general discussion to an area that is not often considered or possibly simply taken for granted when working on spoken languages alone.

4 Iconicity Effects on Phonological Representation

The topic of iconicity in signed languages is vast, covering all linguistic areas – e.g. pragmatics, lexical organization, phonetics, morphology, the evolution of language – but in this chapter only aspects of iconicity that are specifically relevant to phonological and morphophonemic representation will be discussed in depth. This treatment of iconicity presented here is fresh because research on phonology and research concerning iconicity have been taken up by sub-fields completely independent from one other, one side sometimes even going so far as to deny the importance of the other side. Iconicity has been a serious topic of study in cognitivist, semiotic, and functionist linguistic perspectives, most particularly dealing with productive, metaphoric, and metonymic phenomena (Wilcox 2001; Russo 2005; Taub 2001; Sallandre and Cuxac 2007). In contrast, with some notable exceptions, phonology has been studied within a generative approach, using tools that make as little reference to meaning or iconicity as possible.

"Iconicity" refers to the mapping of a concrete source domain and the linguistic form (Taub 2001); it is one of three Peircean notions of iconicity, indexicality and symbolicity (Peirce 1931[1958]). From the very beginning iconicity has been a major topic of study in sign language research. It is always the "800-lb. gorilla in the room," despite the fact that the phonology can be constructed without it. Stokoe (1960), Battison (1978), Friedman (1976), Klima and Bellugi (1979), Boyes Braem (1981), Sandler (1989), Brentari (1998) and hosts of references cited therein have all established that ASL has a phonological level of representation using exclusively linguistic evidence based on the distribution of forms – examples come from slips of the tongue, minimal pairs, phonological operations, and processes word-formation (see Leuninger *et al.* 2007). In native signers, iconicity has been shown experimentally to play little role in first-language acquisition (Bonvillian 1990) or in language processing in native signers; Poizner, Bellugi, and Tweeney (1981) demonstrated that iconicity has no reliable effect on short-term recall of signs; Emmorey *et al.* (2004) showed specifically that motor-iconicity of signed languages (involving movement) does not alter the neural systems underlying tool and action naming. Thompson, Emmorey, and Golan (2005) have used "tip of the finger" phenomena (i.e. almost – but not quite – being able to recall a sign) to show that the meaning and form of signs are accessed independently, just as they are in spoken languages. Yet iconicity is always there, and every one of these authors mentioned above also acknowledges that iconicity is pervasive.

Iconicity has been dealt with in relative, rather than absolute terms. Frishberg (1975) and Klima and Bellugi (1979) have established that signed languages become "less iconic" over time, but iconicity never reduces to zero and continues to be productive in contemporary signed languages; however, there is no means to quantitatively and absolutely measure just how much iconicity there is in a sign language lexicon. The question, "Iconic to whom, and under what conditions?" is always relevant, so we need to acknowledge that iconicity is generation-specific (signs for TELEPHONE have changed over time, yet both are iconic),

context-specific (the sign for PERIPHERAL is different for a part of the city and for a part of a computer system, yet both are iconic), and language-specific (signs for TREE are different in Danish, Hong Kong, and American Sign Languages, yet all are iconic). Except for a restricted set of cases where entire gestures from the surrounding (hearing) community are incorporated in their entirety into a specific sign language, the iconicity resides in the sub-lexical units, either in classes of features that reside at a class node or in individual features themselves.[12] There may be several layers or types of "resemblance" and not all are appropriately called iconicity (Taub 2001; Russo 2005). The resemblance may be a direct one between the form and the source domain; these are termed "iconic", or it may be an extension of this primary connection in another domain ("metaphoric"). It may be part of the etymology of a dictionary entry ("frozen") or it may be latent in the dictionary entry but emerge in particular linguistic contexts ("dynamic"). Finally, iconicity is thought to be one of the factors that makes signed languages look so similar (Guerra 1999; Guerra *et al.* 2002; Wilcox 2010; Wilbur 2010), and sensitivity to and productive use of dynamic iconicity may be one of the reasons why signers from different language families can communicate with each other so readily after so little time, despite crosslinguistic differences in lexicon and in many instances, also in grammar (Russo 2005). Learning how to use iconicity productively within the grammar is undoubtedly a part of acquiring a sign language.

I will argue that iconicity and phonology are not incompatible. Phonology is both an inventory and the distribution of its elements. Now, after all the work over the last several decades showing indisputably that signed languages have phonology and duality of patterning, one can only conclude it is the *distribution* (and not the necessarily the inventory) that must be arbitrary and systematic in order for phonology to exist. Iconicity should not be thought of as either a hindrance or as in opposition to a phonological grammar, but rather another mechanism, on a par with ease of production or ease of perception, that contributes to inventories. Saussure was not wrong, but since he based his generalizations on spoken languages, his conclusions are based on tendencies in a communication modality that can only use iconicity on a more limited basis than signed languages.

Iconicity contributes to the phonological shape of forms more in signed than in spoken languages, so much so that we cannot afford to ignore it. I will show that iconicity is a strong initial factor in building signed words, but it is also restricted in outputs, and it can ultimately give rise to arbitrary distribution in the morphology and phonology. Furthermore, I will explain what phonology and phonetics contribute to this interaction, each in its own manner.

4.1 General Effects

It has been shown that the phonological grammar of a sign language can be constructed without the use of iconicity, but what problems can be confronted or insights gained from considering it? It would be odd, even counter-productive, not to use iconicity when it is so readily available. It has been said that signed

languages use iconicity "because they can," since the physical properties of entities, as well as their positions and movements can be quite well represented using a visual-gestural communication modality. For this reason, Mary Brennan (2005) proposed that spoken languages use iconicity in a limited way, not because there is a linguistic restriction against it to which sound symbolism and onomatopoetic forms are the exception, but simply because the sound-speech modality is not suited to it. A lexicon using a sound-based medium simply cannot be constructed based on how entities in the world sound.[13]

Let us consider the two contexts in which signed languages arise. In most signing communities of the Deaf World, signed languages are passed down from generation to generation not through families, but through communities such as, schools, athletic associations, social clubs, etc. But initially, before there is a community, per se, signs begin to be used through interactions among individuals – either among deaf and hearing individuals ("homesign systems"), or in stable communities in which there is a high incidence deafness. In inventing a homesign system, isolated individuals live within a hearing family or community and devise a method for communicating through gestures that become systematic (Goldin-Meadow 2003). Something similar happens on a larger scale in systems that develop in communities with a high incidence of deafness due to genetic factors, such as what happened on the island of Martha's Vineyard in the seventeenth century (Groce 1985) and in the case of Al-Sayyid Bedouin Sign Language (ABSL; Sandler *et al.* 2005; Meir *et al.* 2007; Padden *et al.* 2010). In both cases, these systems developed at first within a context where being transparent is important in making oneself understood.

Mapping this path from homesign to sign language has become an important research topic since it allows linguists the opportunity to follow the diachronic path of a sign language *al vivo* in a way that is no longer possible for spoken languages. In the case of a pidgin, a group of isolated deaf individuals are brought together to a school for the deaf. Each individual brings to the school a homesign system that, along with other homesign systems, undergoes pidginization and ultimately creolization. This has happened in the development of Nicaraguan Sign Language (NSL; Kegl *et al.* 1999; Senghas and Coppola 2001). This work to date has largely focused on morphology and syntax, but *when and how does phonology arise in these systems*? Aronoff *et al.* (2008) have claimed that ABSL, while highly iconic, still has no duality of patterning even though it is about 75 years old. It is well known, however, that in first-language acquisition of spoken languages, infants are statistical learners and phonology is one of the first components to appear (Locke 1995; Aslin *et al.* 1998; Creel *et al.* 2004; Jusczyk *et al.* 1993, 1999).

Confusion between the concepts of 'transparency' and 'iconicity' often cloud this discussion. Homesigners must be transparent to be understood, but signed languages are not transparent; that is, nonsigners cannot guess the meaning of current ASL signs. In one of the first studies of transparency (Klima and Bellugi 1979: 22), 10 hearing participants were shown 90 signs of abstract and concrete nominals, and were reported to be able to make reasonable guesses about meaning for only nine of the 90 signs (10%). Even when the task was multiple-choice,

the participants could not provide correct answers at a level above chance; therefore, we see from Klima and Bellugi's experiment that ASL is not transparent.

In the next sections we will see examples of iconicity and arbitrariness working in parallel to build words and expressions in signed languages, using the feature classes of movement, orientation, and handshape. The morphophonology of word formation exploits and restricts iconicity at the same time; it is used to build signed words, yet outputs are still very much restricted by the phonological grammar. To make the point that sign languages are much more than their iconic elements, more attention is paid to the small set of forms in sign languages that contradict iconicity (e.g. the sign SLOW signed in its emphatic form surfaces with a faster movement than the base form: Klima and Bellugi 1979), but the more pervasive phenomenon is the use of iconicity and phonology together. The examples below show how iconicity and phonology can be disentangled from one another when both are present.[14]

4.2 Directional Path Movement and Verb Agreement

One area in sign language grammars where iconicity plays an important role is verb agreement. Salience and stability among arguments may be encoded not only in syntactic terms, but also by visual-spatial means. Moreover, path movements, which are an integral part of these expressions, have a phonological location in the feature tree. Tracing the trajectory from a homesign to a pidgin to a creole system of such forms is spelled out in Senghas (1995) for Nicaraguan Sign Language, and for ABSL (Meir *et al.* 2007; Padden *et al.* 2010). Here the phonological consequences of this iconic ability will be discussed.

There are three types of verbs attested in signed languages (Padden 1983): those that do not manifest agreement ("plain" verbs), and those that do, which divide further into those known as "spatial" verbs, which take only source-goal agreement, and "agreement" verbs, which take both source-goal agreement, as well as object and potentially subject agreement (Meir 2002; Meir *et al.* 2007). Meir argues that the main difference between verb agreement in spoken languages and signed languages is that verb agreement in signed languages seems to be thematically, rather than syntactically, determined. Agreement typically involves the representation of phi-features of the NP arguments, and functionally it is a part of the referential system of a language. Typically in spoken languages there is a closer relationship between agreement markers and structural positions in the syntax than between agreement markers and semantic roles; however, sign language verbs can agree not only with themes and agents, but can also agree with their source and goal arguments (Kegl 1985 was the first to note this). The combination of syntactic and semantic motivations for agreement in signed languages was formalized as the "direction of transfer principle" (Brentari 1988), but the analysis of verb agreement as having an iconic source was first proposed in Meir (2002).

Agreement verbs manifest the transfer of entities, either abstract or concrete. The locational-loci of sign language verb agreement are regarded as visual

manifestations, overt indices of the pronominal elements in question, rather than of grammatical categories such as gender or number (cf. Meir 2002 and references cited therein). Crucially, Meir argues that "DIR," which is an abstract construct used in a transfer (or directional) verb, is the iconic representation of the semantic notion "path" used in theoretical frameworks, such as Jackendoff (1996: 320); DIR denotes spatial relations. It can appear as an independent verb or as an affix to other verbs. This type of iconicity is rooted in the fact that referents in a signed discourse are tracked both syntactically and visuo-spatially; however, this iconicity is constrained by the phonology. Independently a [direction] feature has been argued for in the phonology, indicating a path moving to or from a particular plane of articulation (Brentari 1988, 1990a, 1998).[15]

The abstract morpheme DIR and the phonological feature [direction] are distributed arbitrarily both across sign languages (Mathur and Rathmann 2010) and language-internally. In ASL it can surface in the path of the verb or in the orientation; that is, on one or both of these parameters. It is the phonology of the stem that accounts for the distribution of orientation or path as markers in the set of agreement verbs in ASL predicting whether it will surface, and if so, where it will surface. Figure 21.7 provides examples of how the morphological and phonological structures work together, along with diagrams of the [direction] feature involved. In Figure 21.7a we see an example of the agreement verb, APOLOGIZE, that takes neither orientation nor source-goal properties; signs in this set have been argued to have eye gaze agreement with the object (Bahan 1996).[16] The phonological factor relevant here is that many signs in this set have a distinct place of articulation that is on or near the body. In Figure 21.7b we see an example of an agreement verb that takes only the orientation marker of agreement, SAY-YES; this verb has no path movement in the stem that can be modified in its beginning and ending points (Askins and Perlmutter 1995), but the affixal DIR morpheme is realized on the orientation, palm facing a locus is the object plane.[17] In Figure 21.7c we have an example of an agreement verb that has a path movement in the stem

	a. APOLOGIZE	b. SAY–YES	c. HELP	d. PAY
Path marker	ø	ø	+	+
Orientation marker	ø	+	ø	+
Direction feature	ø	To object	From subject to object	To object

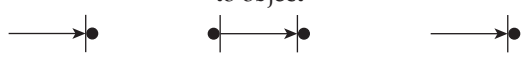


Figure 21.7 Examples of verb agreement in ASL.

– HELP – whose beginning and endpoints can be modified according the subject and object locus. Because of the angle of wrist and forearm, it would be very difficult (if not impossible) to modify the orientation of this sign (Mathur and Rathmann 2006).[18] In Figure 21.7d we see an example of the agreement verb PAY that expresses the DIR verb agreement on both path movement and orientation; the path moves from the payer to the payee, and the orientation of the fingertip is towards the payee at the end of the sign. The analysis of this variation depends in part on the lexical specification of the stem – whether orientation or path is specified in the stem of the verb or supplied by the verb-agreement morphology (Askins and Perlmutter 1995) – and in part on the phonetic-motoric constraints on the articulators involved in articulating the stem – i.e. the joints of the arms and hands (Mathur and Rathmann 2006).

This analysis of the expression shows that the iconic elements and phonological elements of verb agreement are distributed in an arbitrary, but systematic, way that can only be captured by an account of the relevant phonological features of handshape, orientation, movement, and place of articulation (including the appropriate planes of articulation), and constraints on their phonetic realization in the allomorphy of agreement morphology.

4.3 Movement in Event Structure

Wilbur has been involved in research on the relationship between movement and meaning in signed languages since the 1980s (Wilbur *et al.* 1983, 1999, 2008). Her work has analyzed the prosodic uses of movement for stress, accent, and emphasis as well as the use of movement in aspectual morphology. She recently developed the Event Visibility Hypothesis (Wilbur 2008, 2010), which is a proposal for how the structure of predicates in signed languages adheres to a type of mapping between event structure and phonological form.

In English, event structures are inaccessible via the phonology, although these structures are recoverable through syntactic tests; Wilbur argues that in signed languages, event structure is expressed in the phonology: States (Ss) are [–dynamic] and have no movement; processes (Ps) are [+dynamic] and have a movement; telic events transitions between two non-identical sub-events (P->S), and achievements are transitions between two non-identical sub-events (S->S). Furthermore, processes are homogenous and exhibit no changes other than the passage of time, while telic and inchoative events are heterogeneous. Brentari's phonological movement inventory (1998) is correlated with the event structure of categories of predicate signs, which are grouped into those that are atelic (states and processes), telic punctual transitions, and telic non-punctual transitions. The semantic, syntactic tests cannot be reiterated in the space allotted here, but evidence is provided in Wilbur (2010) that these distinctions are part of the semantic *Aktionsart* of the event.

Why is phonology a part of this abstract semantic analysis of event structure? Wilbur argues that by using the features and feature geometry proposed in Brentari (1998) it can be shown that the morphophonology of sign language predicates

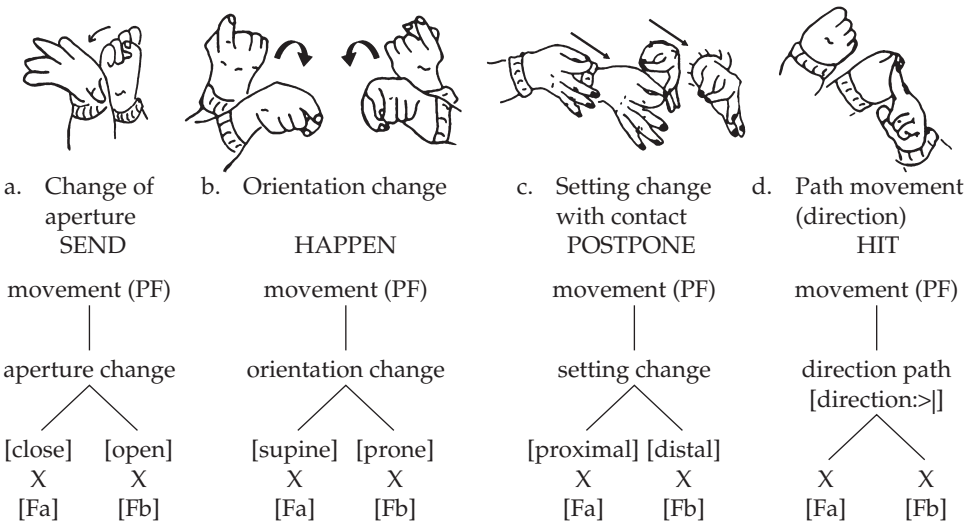


Figure 21.8 Examples of telic procedures according to the Event Visibility Hypothesis (Wilbur 2006).

reflects the temporal components of Pustejovsky's (1995) types of events. These are states, processes, and transitions, which include achievements, and accomplishments). Wilbur further claims that aspectually modified forms, such as resultative, incessant, and continuative, are compositional (each piece contributes meaning), dividing up events into initial, internal, and final temporal sub-periods. Examples include movements for telic and atelic predicates, which exploit changes in the movement features and the associated segmental structure proposed in the Prosodic Model. The features of movement (the prosodic features shown in Figure 21.2d) are used for this analysis. Telic predicates exploit the transitions between the two different specifications for handshape, orientation, setting or [direction] path movements (examples of telic verbs are given in Figure 21.8a–d; this feature was also used in the analysis of verb agreement in the previous section). The feature matrixes of the two segments are not identical in this case. Atelic predicates contain a [tracing] or [trilled] movement, which corresponds to phonological shape and manner features in the Prosodic model. Crucially, the segments are identical; there is no change in the feature matrices of the two X slots, only extension in time (examples given in Figure 21.9a–d). Each of the verbs in Figure 21.9 has a [tracing movement], which specifies the shape, and a [trilled] manner feature that indicates that the movement is repeated an uncountable number of times.

This analysis has been able to establish that there is a relationship between event structure and phonology (meaning and form) that is both iconically motivated and phonologically constrained, and Wilbur has argued that these structural components of predicates are one reason why sign languages look so similar to one another.

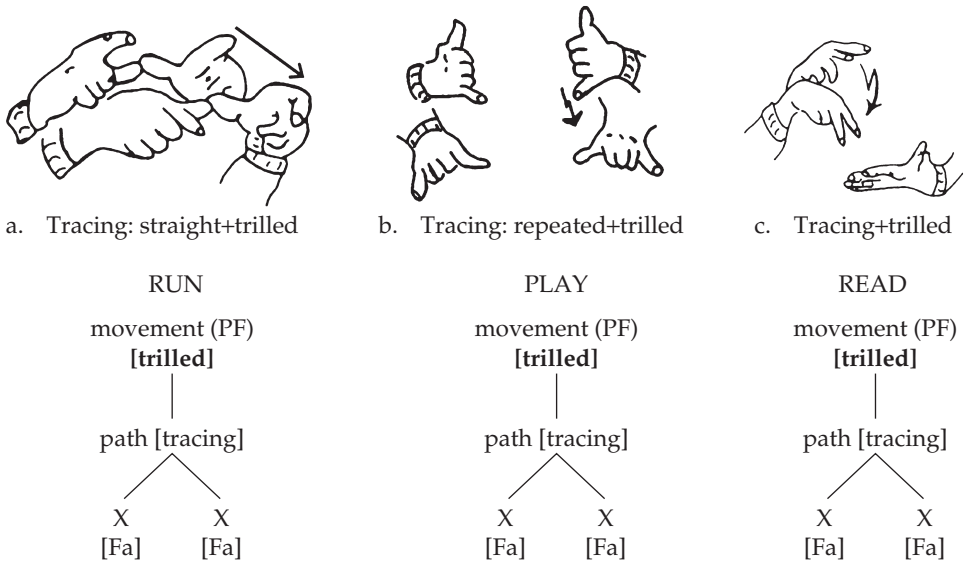


Figure 21.9 Examples of atelic predicates according the Event Visibility Hypothesis (Wilbur 2006).

4.4 Orientation in Handshape in Classifier Constructions is Arbitrarily Distributed

An additional iconic source for a structure that is ultimately distributed arbitrarily involves the orientation of the hand in the handshape of classifier constructions. For our purposes here, classifier constructions can be defined as complex predicates in which movement, handshape, and location are meaningful elements; we focus here on handshape. We will use Engberg-Pederson's (1993) system, given in (7), which divides the classifier handshapes into four groups. Examples of each are given in Figure 21.10.

(7) Categories of handshape in classifier constructions (Engberg-Pedersen 1993)

a. *Whole entity*: these handshapes refer to whole objects (e.g. "1-HANDSHAPE: person" (Figure 21.10a))
b. *Surface*: these handshapes refer to the physical properties of an object (e.g. "B-B-handshape: flat_surface" (Figure 21.10b)
c. *Limb/body part*: these handshapes refer to the limbs/body parts of an agent (e.g. V-handshape: by_legs (Figure 21.10c)[19]
d. *Handling*: these handshapes refer to how an object is handled or manipulated (e.g. "S-handshape: grasp_gear_shift" (Figure 21.10d))

Benedicto and Brentari (2004) and Brentari (2005) argued that, while all types of classifier constructions use handshape morphologically in a general way, only handshapes in classifier constructions of the handling and limb/body part type

a. Whole entity (1-HS: person). b. Surface extension (B-B-HS: flat surface).

Phonological use of orientation.

c. Body part (1V-HS: person). d. Handling (S-HS: grasp).

Morphological use of orientation.

Figure 21.10 Examples of the distribution of phonological and morphological use of orientation in classifier predicates. Whole Entity and Surface/Extension classifier handshapes do not allow morphological use of orientation (top) while Body Part and Handling classifier handshapes do allow morphological use of orientation (bottom).

can use *orientation* in a morphological way, while whole entity and surface cannot. This is shown in Figure 21.10, which illustrates the variation of the forms using orientation phonologically and morphologically. While the whole entity classifier in Figure 21.10a "person_upside_down" and the surface classifier in Figure 21.10b "flat_surface_upside_down" are not grammatical if the orientation is changed (indicated by an "x" through the ungrammatical forms).[20] The body part classifier in Figure 21.10c "by-legs_be_located_upside_down" and the handling classifier in Figure 21.10d "grasp_gear_shift_from_below" are grammatical when articulated with different orientations.

This analysis requires phonology because the representation of handshape must allow for subclasses of features to function independently with respect to the phonology and morphology of the language according to the type of classifier being used. In all four types of classifiers, part of the orientation specification expresses a relevant handpart's orientation (palm, fingertips, back of hand, etc.) toward a place of articulation, but only in body part and handling classifiers is it allowed to function morphologically as well as phonologically. It has been shown that these four types of classifiers have different syntactic properties as well (Benedicto and Brentari 2004; Grose *et al.* 2007).

It would certainly be more iconic to have the orientation expressed uniformly across the different classifier types, but the grammar does not allow this. We therefore have evidence that iconicity is present but constrained in the use of orientation in classifier predicates.

4.5 Conventionalization

A final example of the intertwined nature of iconicity and phonology addresses *how* a phonological distribution might emerge in sign languages over time (Brentari *et al.* forthcoming).[21] Productive handshapes were studied in adult native signers, hearing gesturers (without using their voices), and homesigners in handshapes – particularly, the selected finger features of handshape. Selected fingers indicate which fingers are active in the handshape. The results show that the distribution of selected finger properties is reorganized over time.

We classified handshapes into three levels of selected finger complexity. *Low complexity* handshapes have the simplest phonological representation (Brentari 1998), are the most frequent handshapes crosslinguistically (Hara 2003; Eccarius and Brentari 2007), and are the earliest handshapes acquired by native signers (Boyes Braem 1981). *Medium complexity* and *High complexity* handshapes are defined in structural terms – i.e. the simpler the structure the less complexity it contains. *Medium complexity* handshapes include one additional elaboration of the representation of a [one]-finger handshape, either by adding a branching structure or an extra association line. *High complexity* handshapes included all other handshapes. Examples of low and medium complexity handshapes are shown in Figure 21.11.

The selected finger complexity of two types of productive handshapes was analyzed: those representing *objects* and those representing the *handling* of objects (corresponding to whole entity and handling handshapes in a sign language). The pattern that appeared in signers and homesigners showed no significant differences: relatively higher finger complexity in object handshapes and lower for handling handshapes (Figure 21.12). The opposite pattern appeared in gesturers, which differed significantly from the other two groups: higher finger complexity in

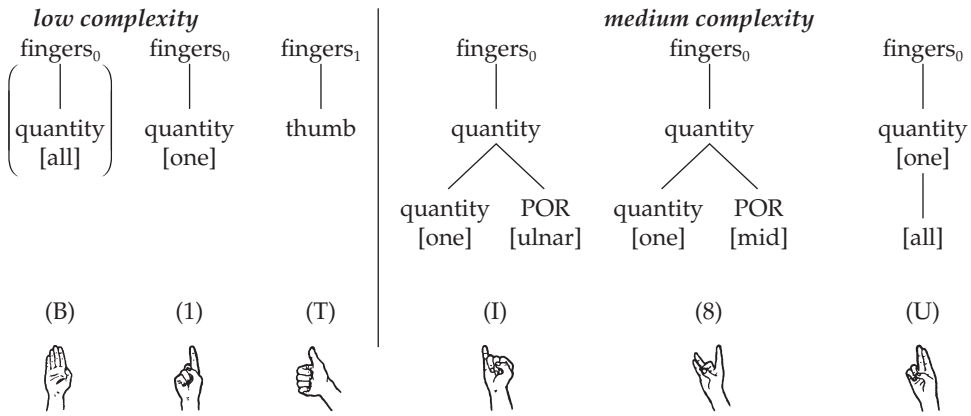


Figure 21.11 The three handshapes with low finger complexity and examples of handshapes with medium finger complexity. The parentheses around the B-handshape indicate that it is the default handshape in the system.

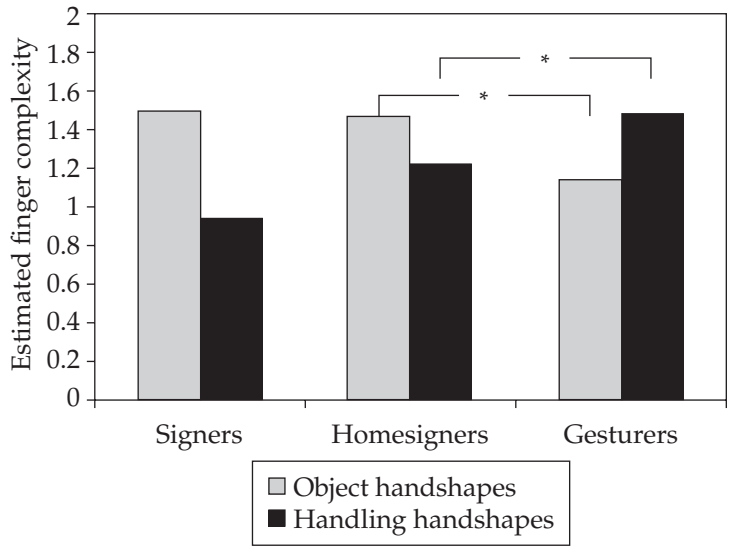


Figure 21.12 Estimated mean finger complexity for *Object* handshapes and *Handling* handshapes in signers, homesigners, and gesturers.

handling handshapes and lower in object handshapes. These results indicate that as handshape moves from gesture to homesign and ultimately to a sign language, object handshapes *gain* finger complexity and handling handshapes *lose* it relative to their distribution in gesture. In other words, even though all of these handshapes are iconic in all three groups, the features involved in selected fingers are heavily reorganized in sign languages, and the homesigners already display signs of this reorganization.

To summarize this section on iconicity, one can say that each of the elements discussed is iconic and, crucially, also phonological. It has been observed that the co-speech gestures of speakers during narration (see McNeill 2005) contain some of the same surface elements of movement, orientation, and handshape; however, phonology emerges from these properties only when these elements become reorganized (or conventionalized) and assume an arbitrary distribution. Iconicity, like ease of articulation and ease of perception, is a factor that contributes to the phonological inventories of sign languages, and based on the work presented in this section, the distribution of the material is more important for establishing the phonology of signed languages than the source of that material – iconic or otherwise.

5 Conclusion

The more phonologists focus on the physical manifestations of the system – the vocal tract, the hands, the ear, the eyes – signed and spoken language phonology

will look different. The more focus there is on the mind/brain the more sign language and spoken language phonologies will look the same.

This chapter was written in part to answer the following questions: "Why should phonologists, who above all else are fascinated with the way things *sound*, care about systems without sound?" How does it relate to their interests?" The short answer to those questions is that that by looking at the differences and similarities in signed and spoken languages, aspects of work on spoken languages can be seen in a surprising new light, because the range of possibility in expression is considerably broadened, as we see, for example, in the work on visual prosody. Phonologists are in a privileged place to see these differences, because, unlike the case of semantics or syntax, the language medium affects the organization of the phonological system. Using work on signed languages, phonologists can broaden the scope of the discipline to include issues of modality and iconicity, thereby acknowledging that phonology has the potential to exploit a greater range of expressive power than previously thought.

ACKNOWLEDGMENTS

Portions of this chapter have appeared in previous works: Parts of Section 2 and 3.2 are discussed in Brentari (2007). The arguments of Section 3.3 are taken from Benedicto and Brentari (2004) and Brentari (2005). Section 4.3 is drawn from Brentari (2006) and Brentari and Brentari *et al.* (2011). Portions of Section 4.4 appeared in Brentari (2002).

NOTES

1 The terms Handshape (HS), Place of Articulation (POA), and Movement (M) are used for clarity and ease of exposition here. "Handshape" seems to imply just one hand, but there are also signs that use the arm, or both hands, so "Articulator" is the term sometimes used to cover all of these. "Place of Articulation" is sometimes referred to as "Location" in the sign language literature. And in some models there is no "Movement," but rather "Manner." A structure that includes the non-manual behaviors of the face and body has not been fully worked out, so these are not included in Figure 21.1.

2 I refer the reader to Brentari (1995), Section 1.1, for a history and description of the featural content of the classes of features, known as "parameters" in sign language phonology. Figure 21.1 is a representation from the Prosodic Model (Brentari 1998), which has a particular organization of features and segments, but the point here is only that the root is a lexeme, rather than a C- or V-unit, and this is common to models by Sander (1989) and van der Hulst (1995, 2000), and Channon (2002) as well as the Prosodic Model.

3 See also Jantunen and Takkinen (2010) and Brentari (1995, Section 2.1.2) for more background on the sign language syllable.

4 There is some variation in the constituent claimed to be associated with this cue. It has been documented in Swiss German Sign Language as the P-Phrase (Boyes Braem 1999), and in ASL as the I-Phrase (Wilbur and Patschke 1998).
5 The dominant hand is the hand used for one-handed signs, and it is the hand on which the more complex handshape appears in two-handed signs. The less complex hand in a two-handed sign is called the nondominant hand (Battison 1978).
6 Not all aspects of sounds are processed cortically; pitch appears to be detected at the level of the brain stem, which is physiologically peripheral with respect to the cortex (Xu *et al.* 2006).
7 Rhythmic cues are not used at the word level in ASL; they begin to be in evidence at the phrasal level (Miller 1996).
8 These constraints hold for lexemes, but they may be violated in ASL compounds.
9 This point has been addressed similarly in van der Hulst (2000).
10 These are surface representations in English and ASL. In English the /u/ in /dud/ is lengthened before a voiced coda consonant, resulting in an output [du:d].
11 [Intensive] forms have a geminated first segment – GOOD vs. GOOD [intensive] 'very good', LATE vs. LATE [intensive] 'very late', etc. – but no lexical contrast is achieved by segment length.
12 Examples of such gestures that are co-opted "whole-cloth" in ASL are PRAY and SO-SO; in Italian Sign Language (LIS) PERFETTO "perfect," SOLDI "money," and COME/PERCHÉ "why." No more will be said about this type of grammaticalization. We leave these cases aside; these transparent forms are rare.
13 Iconicity does exist in spoken languages in reduplication (e.g. Haiman 1980) as well as expressives/ideophones. See, for example, Bodomo (2006) for a discussion of these in Dagaare, a Gur language of West Africa. See also Okrent (2002), Shintel *et al.* (2006); Shintel and Nussbaum (2007) for the use of vocal quality, such as length and pitch, in an iconic manner.
14 See also van der Kooij (2002) and Eccarius (2008) for discussions of the interaction between phonology and iconicity.
15 The vertical is one of the three possible planes of articulation in signing space – horizontal, vertical, or midsagital (Brentari 1998).
16 There is debate about exactly what role eye gaze plays in the agreement system, but that it plays a role is not controversial (Neidle *et al.* 2000; Thompson and Emmorey 2006).
17 There is an emphatic form that has a path movement, added but this is not the form typically used.
18 In many other sign languages, the verb HELP is articulated differently, and in Italian Sign Language, German Sign Language, and Israeli Sign Language HELP has both orientation and path markers for agreement.
19 In ASL we have found that the V-handshape "by-legs" can function as either a body or whole entity classifier.
20 Orientation differences in whole entity classifiers are shown by signing the basic form, and then sequentially adding a movement to that form to indicate a change in orientation.
21 This work is being carried out thanks to NSF grants BCS 0112391 and BCS 0547554 to Brentari, and is being carried out collaboratively with Susan Goldin-Meadow, Marie Coppola, Laura Mazzoni, along with consultants Roberto Ajello and Virginia Volterra.

22 Language Games

BERT VAUX

1 Introduction

Language games (or ludlings) which manipulate the phonological structure of words have been recruited as evidence in phonological debates since the beginnings of generative phonology in the 1960s, and have been of general interest to linguists since at least Sapir (1915) and Jespersen (1922). Bruce Bagemihl, whose 1988 and 1995 surveys I aim to complement here, identifies three types of work on language games: (i) descriptions of individual language games; (ii) use of games to illustrate workings of individual languages; and (iii) use of games to illustrate workings of human language as a whole (1995: 700).

Though there is some debate about the reliability of language games for the latter two purposes (see e.g. Churma 1985; Zwicky and Pullum 1987; Bertinetto 1992; 2004; Rizzolo 2007), phonologists generally agree on their value as external evidence Kenstowicz and Kisseberth (1979); Ohala (1986); Bagemihl (1995: 697) who states that language games are quantitatively but not qualitatively different from natural language processes, Derwing and de Almeida (2009), and Haspelmath (2004). Researchers have employed language games to argue for elements of representations such as phonemes, underlying representations (Botne and Davis 2000), syllable structure (Sherzer 1970; Treiman 1983; Tateishi 1989; Davis and Hammond 1995), psychological reality of rules (Campbell 1980; Guimarães and Nevins 2006), underspecification and prespecification (Kaun and Harrison 2001), contour segments (Cowan and Leavitt 1982), feet (Gil 2002), tonal melodies (Hombert 1973, 1986; Ịṣọla 1982; Campbell 1986), and roots and templatic morphology (al-Mozainy 1982; Heath 1987; Prunet, Beland, and Idrissi 2000), as well as more general characteristics of the phonological component such as inventories and phonotactics (Esper 1925; Diehl and Kolodzey 1981; Campbell 1986), autosegmentalism (Clements 1985; Vago

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

1985, 1988), prosody and mapping (McCarthy 1991; Plénat 1995), and output-output faithfulness constraints (Tessier 2006b).

Work on language games since the early 1990s has moved away from this representational focus in at least three respects. First, it has shifted its focus from subtleties of phonological representation to larger cognitive issues of operations, acquisition, and architecture such as opacity, underdetermination, movement, copying, and the encoding of linear precedence.

Second, following Ohala's 1986 identification of invented language games as a prime source of external evidence in phonology, since extant language games may involve ritualized conventions whereas invented language games can be controlled experimentally and force participants to construct generalizations from scratch, language game research has moved in a more experimental direction, extending to include the acquisition, evaluation, and production of artificial language games and toy grammars. In this category we include acquisition and manipulation of novel language games for existing languages (Treiman 1983; Derwing, Dow, and Nearey 1988; Bertinetto 1988, 1992; Pierrehumbert and Nair 1995; Derwing and de Almeida 2009), cryptophasias or secret sibling languages (Luria and Yudovich 1959; Zazzo 1960; Diehl and Kolodzey 1981; Malmstrom and Silva 1986; Mogford 1993), and artificial language-learning experiments (Gomez and Gerken 2000; Guest, Dell, and Cole 2000; Pater and Tessier 2003; Pycha *et al.* 2003; Wilson 2003; Koo and Cole 2006, Peperkamp and Dupoux 2007).[1] Within this latter category a range of methodologies have been employed in addition to the usual production tasks, including testing generalization, from impoverished data (Wilson 2006), concept generalization (Treiman 1983; Jaeger 1986), embedding artificial languages inside fairy tales (Peperkamp's "accented French"); assessing the grammaticality of forms in a toy language (or assessing the likelihood of a form belonging to the toy language in question; Pycha *et al.* 2003; Nevins and Endress 2007); testing the learnability of different game types (Esper 1925; Treiman 1983); evaluating the recallability of different game types (Kiparsky and Menn 1977); and repetition latency (Onishi, Chambers, and Fisher 2002, Chambers, Onishi, and Fisher 2003; Koo and Cole 2007) and accuracy (Brown and Hildum 1956; Gathercole, Willis, and Baddeley 1991).

Third, the advent of Optimality Theory in 1993 has highlighted a range of questions for which language games are directly relevant, particularly regarding the question of innate biases applied to language games. In OT, all human grammars consist of variant rankings of a universal set of constraints. If this is so, how do the seemingly parochial, unnatural, and highly language-specific processes found in language games, such as Spaka optional n-insertion (Diehl and Kolodzey 1981: 422) fit into this picture? Are grammars monostratal, and if so, how is the apparent ability of language games to target different levels of representation (see Section 2.3.1) to be explained? If learning is deterministic, as Tesar and Smolensky (2000) and other leading optimality-theoretic acquisition models propose, how is one to account for the enormous range of variation observed in the set of generalizations hypothesized by learners upon exposure to a consistent data set? Finally, one of the most intriguing issues in language game research is the appearance of

spontaneous opacity and ineffability effects (see Section 2.1.2.2), for which learners have no overt evidence in the training data, and which are predicted not to arise by current optimality-theoretic models of acquisition.

In this chapter I report on recent research that seeks to investigate questions such as these experimentally, and thereby to situate our understanding of the phonological component of the language faculty within its larger cognitive psychological context.

2 Issues

As already mentioned, language games are a well-known source of external evidence[2] for independently-posited phonological representations and phonotactics. What I propose to do in this chapter is not to rehash these arguments, but rather to analyze the ludling data in their own right,[3] in terms of their acquisition and what they reveal about spontaneous generalizations and inductive biases. The fact that given ambiguous and incomplete data, people show variation, but this variation appears to be constrained in principled ways and learners sometimes posit orderings, rules, or constraints not found in their native language, is truly interesting and unique to language games as opposed to other sources of evidence about the phonological component of Universal Grammar.

In what follows, I focus first on issues involving acquisition of language games and of phonological systems in general. I concentrate on the central issues of variation (interesting in its own right because it reveals heterogeneity in the precise nature of ludling formulations for games even as well-known as Pig Latin), underdetermination (and in particular the fact that learners' selection from the infinite range of hypotheses compatible with a given data set appears not to be entirely random), and the construction of underlying representations. I then turn to experimental work on language games, artificial languages, and the like, focusing on findings bearing on what I consider to be the key issues in recent research in this area, phonotactics and the encoding of linear sequencing. Finally, I turn to two aspects of language game research of particular importance to the debate between rule- and constraint-based models of phonology: ordering effects and naturalness.

2.1 Acquisition Issues

The question of how a phonological system is acquired lies at the very heart of phonological theory. Language games have a key role to play in this domain, because they enable us to investigate acquisitional questions on humans of all ages in a fairly straightforward manner, they allow for experimental control of the data and the acquisition process in ways that are not possible with natural acquisition, and they make it possible to present learners with significantly more impoverished sets of primary linguistic data than is possible in natural acquisition settings. In this section I discuss three central issues at the nexus between recent

acquisition and language game research: variation, underdetermination and analytic biases, and the construction of underlying representations.

2.1.1 Variation Variation is perhaps the most pervasive characteristic of language games and artificial language learning. For example, Vaux and Nevins' Pig Latin survey of 447 individuals revealed 21 different ways of dealing with vowel-initial words such as enter, summarized in (1).

(1) Pig Latin treatments of *enter*[4]

Form	*Number*	*%*
enter-ay	173	39
enter-yay	65	15
enter-way	61	14
enter-hay	44	10
ter-en-ay	38	9
no output sounds good	36	8
ter-ent-ay	13	3
y-enter-yay	3	1
enter-ent-ay	2	–
er-ent-ay	2	–
h-enter-hay	1	–
en-ay er-tay	1	–
enter wah/wu (= [wə]?)	1	–
en-way ter-ay	1	–
enter-lay	1	–
nter e-gay	1	–
nter e-way	1	–
ter-nay	1	–
nter e-ay	1	–
enter-nay	1	–
none of the above	28	6

The above example appears to involve variation in the ludling rule(s) postulated rather than in the natural grammars of the individual speakers (pace Barlow 2001a).

In addition to the variation in rule formulation documented in (1) and in more detail with regard to anchor points in Section 2.1.2.1, variation along other phonological dimensions is discussed throughout this chapter, including variation in constraint postulation (2.1.2.2), phonotactic variation (2.2.1), the point in the derivation targeted by ludling rules (2.3.1), and consonant and vowel selection for infixation (2.3.2).

What is the source of this variation, which appears to be qualitatively more extensive than in natural phonological systems? One factor that is likely to be involved is the lack of regulation imposed from above on language games. The

same dynamic can be seen with child vocabulary that slips under the radar of schools, dictionaries, and other bastions of orthographic and phonological control: it is prone to extensive variation (witness the phonological variants of the knife-tossing game mumblety peg documented by Vaux and Golder 2003,[5] including mumbledy-peg, mumbly peg, mumbly pegs, mumblely peg, mumble peg, mummety-peg, mumble-the-peg, fumbled peg, and numblety peg).

This lack of regulation maximizes the underdetermination of games such as Pig Latin; typically a child learning Pig Latin for the first time will be told only that the phrase Pig Latin is transformed into ig-pay atin-lay, leaving undetermined the treatment of consonant clusters, vowel-initial words, words with non-initial stress, and so on. If phonological acquisition were deterministic, based on either (in OT terms) the default ranking of CON provided by Universal Grammar or the ranking of CON required to generate the child's native language, we would not predict the poverty of the ludling stimulus to lead to the wide range of variation that we actually find. What seems more likely is that some form of abductive learning is at work, as suggested by the Stanford Child Phonology school (e.g. Macken and Ferguson 1983): given a body of surface data, the learner's Language Acquisition Device employs the primitives provided by UG (which may include features, logical and combinatorial operators, and the like) to generate a set of hypotheses capable of generating those data. It is likely that these competing hypotheses are assigned probabilities based on elements of prior knowledge (e.g. bias towards formally simpler and more predictive hypotheses); the weighting by priors will serve to skew the range of analyses selected by learners in certain directions, while the probabilistic component of the computation is necessary to explain the range of analyses selected.

Language games also reveal another type of variation of interest for phonological theory, which I term hypervariation. This phenomenon refers, in the case of ludlings, to a speaker having significant latitude to choose how to apply one or more (already determined) rules of a game to a given form. Bedouin Hijazi Arabic and Moroccan Arabic, for example, allow free metathesis of root consonants (Prunet, Béland, and Idrissi 2000: 623). By dint of this generalization the Bedouin form kattab 'caused to write' can become battak, takkab, tabbak, kabbat, or bakkat (Al-Mozainy 1981: 86). Similar optionality in choice of root consonants serving as anchors in Arabic hypocoristics has been documented by Davis and Zawaydeh (2001).

We find another sort of hypervariation outside the Semitic family, where some language games allow for a virtually limitless set of transformations as long as these honor the basic requirements of the game. In the Norwegian ludling Smoi, for example, "the overall transformation principle seems to be as general as this: split the Mandal dialect word into syllables and/or phonemes, then rearrange the sequence of syllables or phonemes in any order you like, provided it is pronounceable" (Jahr 2003: 276). Plénat observes a similar polymorphism in Verlan, where any mapping of the segments of a base word to the minimal word template of Verlan is acceptable as long as it honors certain association conventions, including Exhaustivity, Conservation of Weight and Position, and the Line Crossing Constraint (1995: 26). For consonant-initial trisyllabic words with the syllable

sequence 123, for instance, four outcomes are attested: 312, 231, 321, and null output Ø (but not 132 and 213) (Plénat 1995: 6). The same sort of situation can be argued to obtain in Expletive Infixation in English. For example, the expletive can be inserted before any non-initial foot, allowing for two variants in a tripodic word such as Popocatepetl → Popo-fuckin-catepetl ~ Popocate-fuckin-petl (McCarthy 1982: 578); as the number of feet in an English word is in principle unbounded, so is the number of insertion sites.

The importance for phonological theory of the several types of variation documented in this section should not be underestimated. First of all, the hypervariation phenomena just discussed suggest that the phonological component must include a free choice mechanism of some sort. This can be achieved by postulating the existence of a relevant logical operator involving existential quantification in the phonological component of UG, but poses a slight problem for OT, which neither includes such a mechanism nor provides any straightforward way in which it could be incorporated or replicated in the grammar. (Due to the omnipresent nature of OT constraints, other constraints favoring edges, unmarked segments, and the like will always trump the free choice that might otherwise result from constraint underdetermination. In other words, one can have TETU or hypervariation but not both.) Second, the significant variation within and across individuals attested in the range of hypotheses generated on the basis of ludling stimuli suggests that phonological learning involves a non-deterministic abductive algorithm. Finally, the "craziness" of many of the hypotheses entertained (as in Spaka, a sibling game discussed in Section 2.3.2) suggests that the phonological learning algorithm is not limited by considerations of naturalness or markedness to the extent that leading theories might lead us to expect.

2.1.2 Underdetermination and Prior Biases The significant range of variation documented in the previous section can plausibly be attributed in most cases to underdetermination of the primary linguistic data: for a finite data set, many linguistic generalizations will be capable of generating the attested facts, and learners may either vary in the generalization they select (as in the case of Pig Latin) or formulate a generalization that itself underdetermines the outcome for a given input (as in the case of Expletive Infixation). Interestingly, though, the range of attested variation does not appear to be as broad as one would predict if learners were selecting freely from the range of viable hypotheses.

Consider, for example, one of the most cross-linguistically common ludling types, which permutes the order of syllables in a word. What outputs of such a game will typically be available to the learner attempting to construct an analysis of it? For many natural languages – and particularly for children learning those languages and their attendant games – most or all of the game outputs will be based on mono- or bisyllabic words; outputs based on words of three or more syllables will be vanishingly rare, as Plénat (1995) observes for Verlan. The learner presented with such a situation will of course be unable to infer anything from the behavior of monosyllables, and the behavior of bisyllables will be compatible with many possibilities, including those in (2):

(2) some generalizations compatible with the datum $/\sigma_1\ \sigma_2/ \rightarrow [\sigma_2\ \sigma_1]$

a. reverse the sequence of syllables $/\sigma_1 \ldots \sigma_{n-1}\ \sigma_n/ \rightarrow [\sigma_n\ \sigma_{n-1} \ldots \sigma_1]$
b. exchange the first and final syllables $/\sigma_1 \ldots \sigma_f/ \rightarrow [\sigma_f \ldots \sigma_1]$
c. exchange the first and second syllables $/\sigma_1\ \sigma_2/ \rightarrow [\sigma_2\ \sigma_1]$
d. exchange the final and penultimate syllables $/\sigma_{f-1}\ \sigma_f/ \rightarrow [\sigma_f\ \sigma_{f-1}]$
e. exchange every pair of syllables L→R
f. exchange the syllables within every foot
g. exchange the syllables in bisyllabic words but no others
h. exchange the syllables in words containing an even number of syllables, etc.

But are all such variations possible? Apparently not: only about five of the infinite range of possible permutations are attested in language games (Nevins and Endress 2007):

(3) Attested syllable permutation game types

a.	move first σ to end	Fula	pii.roo.wal → roo.wal.pii
b.	move final σ to beginning	Tagalog	ka.ma.tis → tis.ka.ma
c.	transpose σ_1 and σ_2	Marquesan	nu.ku.hi.va → ku.nu.hi.va
d.	transpose final and penultimate σ	Luchazi	ya.mu.nu.kwe → ya.mu.kwe.nu
e.	invert order of all σ	Saramaccan	va.li.si → si.li.va

All five of the patterns in (3) are compatible with the disyllabic pattern $/\sigma_1\ \sigma_2/ \rightarrow [\sigma_2\ \sigma_1]$, but far from all of the generalizations compatible with the bisyllabic pattern appear to be entertained by learners. Is this a systematic gap, or an accidental one?

Nevins and Endress (2007) addressed this question by conducting a pilot experiment in which participants were presented with transpositions of trisyllabic sequences of nonce syllables: 123 → 321 (e.g. ka.le.bo → bo.le.ka).[6] This transformation is compatible with at least the hypotheses in (2). Interestingly, Nevins and Endress found that some test subjects opted for generalization (2a) and others for (2b), but none opted for (2c) or for the option of exchanging every other syllable (i.e. σ_n with σ_{n+2}). They concluded that prior analytic biases constrain the choice of which generalizations compatible with a given data set a learner entertains. This hypothesis is compatible with Gerken and Bollt's (2008) finding that 9-month-old infants can generalize from three disparate tokens of a pattern (the so-called Goldilocks Rule), but – crucially for our purposes – only if the generalization is one allowed in natural languages.

The ways in which the underdetermination problem is dealt with by learners thus far seem sensible and limited. Once one's empirical net is cast further, though, the picture becomes a bit more complicated. For instance, Wilson (2006) investigated the acquisition of velar palatalization (k → s) in an artificial language, and found that test subjects given underdetermined data sometimes extended the

generalization from the mid vowel contexts in their training set (i.e. stimuli showing /k/ → [s] / _e) to a high front vowel context. Interestingly, no test subjects generalized in the opposite direction (from high vowels to mid vowels). Though this accords well with the historical observation that palatalization tends to be triggered first by high front vowels and then may spread to other high or front vowels, it suggests that this pattern may be attributed to biases in the Language Acquisition Device rather than simple acoustic or articulatory patterns that tend to interfere with acquisition and change.

Even more unexpected is the back-copying overapplication that shows up in some speakers of Pig Latin, producing forms such as oven → w-oven-w-ay and enter → y-enter-y-ay (cf. (1)). One can debate whether this pattern results from a constraint requiring base-reduplicant identity or rather from a particular formulation of precedence instructions of the sort Raimy (2000) employs to derive back-copying in reduplication;[7] either way, the fact remains that these Pig Latin speakers have considered and selected a hypothesis outside the inventory that we might otherwise have thought to be generated by the Language Acquisition Device.

So far in this section we have considered general limitations placed by the Language Acquisition Device on the hypothesis space scanned by the learner. In the remainder of 2.1.2 we examine in more detail three specific types of prior knowledge that shape the hypothesis space: anchor points, avoidance constraints, and hidden generalizations.

2.1.2.1 Anchor Points Nevins and Vaux (2003) and Yu (2003, 2007) proposed that processes such as ⋱ infixation and reduplication can target only a restricted set of anchor points: first syllable, first foot, first consonant, first vowel, stressed syllable, and final syllable. In a representational model of precedence such as the one proposed in Raimy (2000), infixation and reduplication can be thought of as precedence-modifying operations that, by hypothesis, can target only these points. Interestingly, as already alluded to in the discussion of Nevins and Endress (2007), language games seem restricted to these positions as well, even when the training data allow for other possibilities. Language games may exhibit variation between speakers, but this variation is constrained by what is possible given the above inventory of anchor points.

Focusing on shm-reduplication, Nevins and Vaux (2003) found that most survey respondents presented with the input "obscene" selected either obscene-obshmene or obscene-shmobscene. By hypothesis this is because some speakers opt for the first syllable as the locus of the fixed segmentism shm-, while others opt for the stressed syllable. These two hypotheses largely converge for the immense inventory of bisyllabic trochaic words in English, but diverge on iambic words such as obscene, leading plausibly to precisely the variation documented by Nevins and Vaux.

A similar type of variation that arises due to ambiguity in the core of the ludling data involves the difference between placing the shm- after the first consonant and placing the shm- before the first vowel. For words with simple initial Onsets these deliver the same output, but for words with complex Onsets in their initial or stressed syllable, such as "breakfast," they diverge. Nevins and Vaux (2003) found

that while the majority of respondents preferred outputs such as breakfast-shmeakfast, where the fixed segmentism targets the initial (or stressed) vowel, some respondents opted for breakfast-shmreakfast instead, targeting the first consonant. Hammond (1990) found similar variation for the Name Game: some speakers turned Claire into Bo-Baire and others into Bo-Blaire. Pierrehumbert and Nair (1995) found the same sort of variation when they trained subjects in a game that involved infixation of a -VC- sequence after the first consonant in CVC monosyllables. When subsequently presented with CCV-initial words, some test subjects inserted the fixed -VC- sequence after the first consonant and others before the first vowel. (Interestingly, some test subjects treated CC- clusters of falling sonority differently than ones of rising sonority, as I discuss later in this section.)

Finally, variation in anchor choice arises with respect to glides in sequences such as union. Davis and Hammond (1995), Barlow (2001), and Yip (2003) have discussed how games such as Pig Latin can provide insights into whether the glide in such sequences is part of the Nucleus or part of the Onset. Nevins and Vaux (2003) found that most speakers opt for union, shmjunion rather than union, shmoonion, and concluded that the variation is due to whether speakers target the first nuclear segment, or the first vowel.

In his 1988 dissertation Bruce Bagemihl, one of the most ardent proponents of ludlings as an object of linguistic study and as a source of information about possible and impossible phonological operations, constructed an extensive typology of attested and non-attested ludlings. Some of Bagemihl's generalizations are listed in (4).

(4) Unattested ludlings (Bagemihl 1988)

a. No ludling reverses the middle two syllables (e.g. bar.go.tu.li → bar.tu.go.li).
b. No ludling moves the final syllable to the arithmetic middle (e.g. bar.go.tu.li.na → bar.go.na.tu.li).
c. No ludling permutes every other segment in a word (e.g. bram.poj → am.brjop).
d. No ludling permutes feet (e.g. bar.go.tu.li → tu.li.bar.go).
e. No ludling permutes subsegmental features (e.g. tom.duk → nob.tug).

Following Bagemihl's insight that "ludlings extend, modify, or exaggerate attested natural language processes," I concur that precedence-modifying ludlings constitute a rich source of information about spontaneous transformations on phonological representations, free of prescriptive influence, and that given the wide variety of ludling processes, what one does not find can be quite suggestive.

2.1.2.2 Avoidance Constraints One of the most interesting aspects of the acquisition of language games, and indeed of the acquisition of languages in general, is the spontaneous appearance of phonological effects that do not appear to be reasonably warranted by the primary linguistic data. This is perhaps most striking when such effects appear in a language game but are not found in the natural language upon which the game is parasitic; in such cases one assumes that the learner has

observed sufficient data in the host language to know that it does not contain the rule or constraint in question. Nevins and Vaux (2003), for instance, documented the existence of [-anterior] dissimilation among speakers of shm-reduplication, whereby speakers replace the fixed segmentism *shm-* with *sm-* in forms such as Ashmont-smashmont, ash-smash, witches-smitches. Feature dissimilation of this sort also occurs in the child J's reduplicative language game, in which the reduplicant always begins with /b/ unless the base does as well, in which case the /b/ of the reduplicant dissimilates to /p/ (Inkelas 2003). Spontaneous anti-identity effects in echo reduplication of this sort are striking because there is no relevant evidence presented to the learners in question.

Avoidance of total base-reduplicant identity in ludlings[8] involving reduplication with fixed segmentism is even more common, and in fact appears to be the norm in such phenomena. Users of English shm-reduplication opt for a wide variety of avoidance strategies with shm-initial words, including replacing *shm-* in the reduplicant with shn-, shl-, shf-, shv-, shml-, sh-, fl-, vl-, shp-, and r- (Nevins and Vaux 2003a). Similar effects can be found in m-reduplication in Armenian (e.g. pətuʁ 'fruit' → pətuʁ mətuʁ 'tutti frutti'; Vaux 1998) and Abkhaz (e.g. tʃɨ-k' 'horse-indefinite' → tʃɨk' mɨk' 'horses and the like'; Bruening 1997). In these languages m-reduplication works essentially like shm-reduplication in English but with m- as the fixed segmentism rather than shm-; the difference is that Abkhaz employs tʃʲ rather than m- with m-initial words, whereas Armenian chooses to change the first vowel of the reduplicant to -u- in such cases. Hindi v- echo formation behaves in a similar fashion, replacing the fixed segment with ʃ only when the base begins with v (Nevins 2005). Turkish intensive adjectival formation is ambiguous between the feature dissimilation type and the total segment replacement type: it replaces the fixed segment p with m only when the corresponding element in the base is already an obstruent (Kelepir 1999).

The Name Game song, recorded by Shirley Ellis in 1965 and studied by Hammond (1995), contains an avoidance strategy that should look quite familiar in light of the phenomena just described, but remains unprecedented in its attempt to explain the avoidance strategy at length in rhymed metrical verse. The regular workings of the game are exemplified by the first verse of the song (5a), and the avoidance strategy in the second verse (5b).

(5) a. verse 1 of The Name Game (Ellis 1965)
Shirley!
Shirley Shirley bo birley, banana fana fo firley,
Fee fie mo mirley, Shirley!

b. verse 2 of The Name Game
But if the first two letters are ever the same,
I drop them both and say the name,
Like: Bob, Bob – drop the B's: Bo ob
For Fred, Fred – drop the F's: Fo red
For Mary, Mary – drop the M's: Mo ary
That's the only rule that is contrary.

We can infer from the description in (5b) that the Name Game's avoidance strategy applies in cases where one of the fixed segments and the name it targets begin with the same segment, presumably because it would otherwise not be clear that these names had actually undergone the mutations required by the game. Applying the transformations implicit in verse 1 to Bob, for example, should yield **Bob Bob bo bob*, where the final string is phonetically ambiguous between <bob> (having replaced the initial <B> of <Bob> with the fixed <b> of the game) and <Bob> (where the base name has remained unchanged). The strategy invoked to avoid this ambiguity is (in descriptive terms) to drop both the fixed segmentism and the identical segment in the base name, yielding *Bob Bob bo ob*.

Spontaneous avoidance may also appear in the Japanese language game Zuuja-go (Itô, Kitagawa, and Mester 1996, but see Sanders 1999 for a competing analysis not involving avoidance). Zuuja-go normally moves the last syllable in a word to the beginning, as in [takɯʃi:] 'taxi' → [ʃi:takɯ], but according to Itô, Kitagawa, and Mester does not produce outputs for forms such as *kurisumasu* 'Christmas'. Itô, Kitagawa, and Mester attribute the ineffability in this case to constraints on foot shape allowed by the ludling; in a derivational framework, one could propose either that forms like *kurisumasu* do not satisfy the structural description of the ludling rule(s) responsible for Zuuja-go, or that they do satisfy it but the output of the rule(s) in such cases runs afoul of an inviolable constraint. No matter what is responsible for the null output effects in Zuuja-go (as well as in *shm*-reduplication, Pig Latin, and many other ludlings), their existence poses a problem for OT, where every contentful input should produce a contentful output of some type (for reasons outlined in Vaux 2008, pace McCarthy and Wolf 2009).

A Korean language game involving iterative infixation of a fixed CV sequence after every vowel in the base word reveals another type of spontaneous OCP avoidance strategy. In this game, described in Sohn (1987), one inserts after each vowel in a word a fixed consonant (typically p, though any consonant can be chosen) followed by a copy of the vowel, and deletes any Coda consonants following it, as in (6). (It is impossible to determine whether CV insertion precedes or follows Coda deletion.)

(6) Korean iterative infixing language game

a.	UR	/k'aŋtsʰoŋ/	'hopping'
b.	CV insertion	k'apaŋtsʰopoŋ	
c.	Coda deletion	k'apatsʰopo	
d.	SR	[k'apatsʰopo]	

The ludling rules reflected in (6b–c) must follow the regular Korean processes of post-unrelease fortition (which, for example, changes /hakkyo/ 'school' to [hak.k'yo]; Kim-Renaud 1986) and syllabification (which, for example, changes /salm-i/ to [sal.mi]), as shown in (7a) and (7b) respectively:

(7) Ordering of ludling relative to fortition and syllabification

a. fortition precedes ludling: /hakkyo/ → [hapak'yopo], not *[hapakyopo]
b. syllabification precedes ludling: /salm-i/ → [sapamipi], not *[sapaipi]

We therefore appear to have the rule ordering illustrated with relevant derivations in (8), with crucial orderings indicated by square brackets:

(8) Korean rule ordering[9]

	a. UR	/hakkyo/	/salm/	/salm-i/
[[	b. syllabification	hak.kyo	salm	sal.mi
[[	c. fortition	hak.k'yo	–	–
	d. CV insertion	hapak.k'yopo	sapalm	sapamipi
	e. Coda deletion	hapak'yopo	sapa	–
	f. SR	[hapak'yopo]	[sapa]	[sapamipi]

Interestingly, when the fixed consonant is identical to one of the consonants in the base form, the ordering of the ludling processes (8d–e) can change. If we insert -k'V- in /hakkyo/, for example, we expect by (8) to obtain *[hak'ak'yok'o], but what we actually get is [hak'akyok'o], where it appears that the ludling has bled fortition. Similarly, if we insert -mV- with a form such as /nim-i/, the ordering in (8) predicts *[ni.mi.mi.mi] but what we actually get is [ni.mi.i.mi], where the ludling appears to have bled word-level syllabification. I return to this case when surveying ordering effects in Section 2.3.1.2.

2.1.2.3 Hidden Generalizations Just as language games sometimes reveal avoidance constraints that do not appear to play any role or even exist otherwise in the language, they can also allow arguably universal phonological effects such as phonotactic dispreferences and markedness effects to arise that are either non-existent or hidden in the regular phonology of the language. Pierrehumbert and Nair (1995: 98ff.) in their Experiment 1 trained English-speaking test subjects in an artificial language game that inserted VC strings into CVC words, yielding C[əC]VC. The sequence big + əɪ, for example, produced [bəɪɪg]. Once the subjects had internalized this pattern, they were asked to extend it to CCV- words using a new infix (-əl, -ət, or -ək), and were found to vary in their treatment of such cases. Of particular interest for our present purposes is that several subjects treated s + stop sequences differently than stop + r sequences,[10] reflecting a pattern found in many languages of the world (see Vaux and Wolfe 2009) but arguably not evidenced in the regular phonology of English. Whether this represents (in terms of Rule-Based Phonology (RBP)) a spontaneous generalization drawing on elements of UG or (in OT terms) the emergence of a hidden constraint ranking cannot be determined from the available data, but merits further investigation.

Moreton, Feng, and Smith (2005) examined error patterns in a language game called "Sounding Out," in which each consonantal segment is supposed to be

followed by a schwa. The errors of interest to them involved participants inserting the schwa before the consonant instead of after it. Moreton, Feng, and Smith found that these errors predominantly occur when the C is a sonorant consonant (reminding one of the English names for the letters of the alphabet, for example, [si:] <c> vs. [ɛm] <m>), on the basis of which they proposed a covert preference for sonorants as Codas rather than Onsets, part of the general *Peak/X and *Onset/X hierarchies of Prince and Smolensky (1993). As they point out, since these hierarchies play no role in the phonology of ordinary English, the game appears to expose a covert ranking.

To summarize what has been presented in this section, the variation documented in Section 2.1.1 can plausibly be attributed to the fact that the ludling outputs observed by the learner significantly underdetermine the underlying process(es) responsible for those outputs, leaving the learner free to choose from a wide range of hypotheses. Learners appear to entertain only a small subset of the possible hypotheses, though, suggesting that the hypothesis space is significantly (but non-uniquely) constrained by the Language Acquisition Device (cf. Bonatti *et al.* 2005; Moreton 2008a). The LAD appears to provide a limited array of primitives, such as anchor points, from which generalizations can be built, and appears moreover to make available constraint types not found in the host language which can be inviolable (as in Zuuja-go) or perhaps even trigger local rule reordering (as in Korean iterative infixation). It would be interesting to investigate whether any of these spontaneously emerging constraints are truly spontaneous – that is, of the sort that must be constructed on the spot – or whether instead all such constraints are plausible candidates for the universal constraint set CON.

2.1.3 The Construction of Underlying Representations: Underspecification and Free Rides

> Many secret languages show that the speakers of a language have access to representations more abstract than the superficial phonetic level. (McCarthy 1991: 11, trans. BV)

In the previous two sections we examined how language games can be used to investigate the ways in which phonological generalizations involving rules and constraints are acquired. In this section we employ similar strategies to elucidate the nature and construction of underlying representations, and the question of whether these representations can be regulated by Morpheme Structure Constraints (MSCs). These topics have recently acquired new relevance due to the postulation by Prince and Smolensky (1993) of Richness of the Base (ROTB, the idea that there are no constraints on the form that underlying representations may take) and Lexicon Optimization (LO, construction of underlying forms using the same set of ranked constraints employed in reverse for selection of output forms). Classic OT (Kager 1999) and RBP (Kenstowicz 1994a; Vaux 1998) make interestingly different predictions in this regard: OT predicts that URs should depart from input SRs only when motivated by surface phonological alternations, whereas RBP

predicts that URs can diverge from SRs in the absence of surface alternations when (i) the derivation includes one or more neutralization rules, or (ii) the language contains relevant Morpheme Structure Constraints or takes a free ride (in the sense of Zwicky 1970) on a rule motivated by alternations elsewhere in the language.[11]

Consider, for example, the case of Lac Simon, which according to Kaye (1979) possesses a devoicing process that neutralizes the underlying voicing contrast in word-initial obstruents. How will the language treat a borrowed or nonce form such as [panan]? Classic OT, armed with ROTB and LO, predicts that speakers will invariably assign it the underlying representation /panan/, whereas RBP allows for the possibility that individual speakers will give the form a free ride (Zwicky 1970) on their initial devoicing rule and assign it the underlying representation /banan/. RBP moreover allows for violations of the Richness of the Base principle commonly referred to as Morpheme Structure Constraints to constrain the range of possible URs. Lac Simon, for instance, possesses according to Kaye (on the basis of loanwords and the structure of the native lexicon) an MSC that bans underlying stem-initial voiceless obstruents; the effect of this constraint with respect to our hypothetical SR [panan] would be to limit speakers' choice of UR to the free ride option [banan], precisely the option that is explicitly ruled out in Classic OT.

Lac Simon thus appears to consistently map word-initial voiceless obstruents in surface forms to voiced obstruents in underlying forms, by dint of an MSC. However, in a hypothetical language Lac Simon-prime that is identical to actual Lac Simon save that it lacks this MSC, RBP predicts that a surface form like [panan] should be mappable to either /panan/ or /banan/, whereas Classic OT allows only /panan/. This variability in UR choice is predicted by RBP in any case where the phonological derivation includes one or more relevant neutralization rules (in the case of Lac Simon, the word-initial obstruent devoicing rule). But do we find variation in UR choice in languages that contain neutralization rules? And do we find URs that take free rides? For the latter the answer is clearly yes, as shown by Hombert (1973) for a syllable-reversing language game in Bakwiri (which undoes predictable but non-alternating Nasalization before prenasalized stops, for example, kõmbà 'to take care' → mbákò, not expected *mbákõ), Kaye (1979) for Lac Simon, McCarthy (2005) for Sanskrit, Colloquial Arabic, Choctaw, Rotuman, and Japanese, Nevins and Vaux (2008) for Turkish, and so on.

The Cuna syllable-reversing game Sorsik Sunmakke may display relevant variation in UR construction, though the interpretation of the relevant facts is disputed. Sherzer (1970) observes that one sometimes finds variation in the Sorsik Sunmakke outputs for Cuna forms, for example, the Cuna surface form [gammai] 'sleeping' becomes maigab for some speakers and maigam for others. Churma (1985) and Bagemihl (1995) assert that these output variants result from individuals inserting the ludling's syllable permuation rule at different points in the phonological derivation, but the facts are equally consistent with an analysis in which some speakers postulate an underlying representation /gammai/ while others take a free ride on the language's nasalization rule and postulate underlying /gabmai/.

(It should be noted that a free ride is involved in either scenario.) By the same token, Sherzer observes that non-alternating [aili] 'mangrove' becomes either liak or liai in Sorsik Sunmakke, the former displaying a free ride on the language's k-vocalization rule.

We have seen so far in this section that language games can be fruitfully brought to bear on questions of variation, constraint, and overapplication in the construction of underlying forms, in each case supporting the predictions of RBP. But what about the central OT tenet of Lexicon Optimization, which allows for departure from surface forms in the presence of phonological alternations but not otherwise? Using a set of language games in Hungarian and Finnish and a game-like reduplication process in Turkish and Tuvan, Kaun, and Harrison (2001) found that non-alternating harmonic segments in harmonic roots actually undergo harmonic alternations when appropriate phonological contexts are created by a language game or fixed-segment reduplication (e.g. Tuvan *idik* 'boot' → idik-adɨk (not *idik-adik), suggesting that they have undergone a free ride on the harmonic generalizations in the language and have been underspecified for the relevant harmonic features in their underlying representations. Similar effects revealed by a language game in Finnish were documented by Campbell (1980).

The available language game evidence thus makes a number of important suggestions with respect to the construction of underlying representations. First, learning is not deterministic: in cases where the solution is underdetermined by the data, we find variation within and across speakers in the analysis they choose. Second, learning is aggressive: speakers frequently overapply phonological generalizations, resulting in unnecessary violations of surface faithfulness. Third, speakers manipulate constraints on underlying representations, not just surface representations.

When these results are combined with those of Sections 2.1.1 and 2.1.2, what emerges is that phonological language games can shed significant light on central questions of acquisition, most notably involving the algorithm(s) employed in constructing phonological generalizations (be they rules or constraints) and underlying representations. The results of investigations in this domain to date suggest that the hypotheses entertained by learners are both aggressive and highly variable, but that both of these properties are constrained by limitations built into the Language Acquisition Device concerning what phonological elements and logical operators exist and how they can combine.

2.2 Experimental Evidence

Having surveyed the relevance of language games to the study of phonological acquisition, I would now like to turn to the experimental investigation of language games in general and artificial grammars in particular, which has moved to the forefront of phonological inquiry in recent years. Though the range of experimentation conducted in this new field is already quite large, I will focus on two areas of investigation that have been particularly popular and revealing, phonotactics and precedence relations.

2.2.1 Phonotactics Language games can sometimes be used, when naturally-occurring data are unhelpful, to determine whether a phonotactic gap is accidental or systematic (cf. Ohala 1986). For instance, the Lax Vowel Constraint in Standard American English, which according to Cebrian (2002) disallows lax vowels in open stressed syllables (I prefer to formulate the generalization as a ban on the non-low vowels {ɪɛʊ} in word-final position), is suggested to be synchronically active in English by the fact that in the English-based sibling language Spaka (discussed in more detail in Section 2.3.2), lax vowels generated from English tense vowels are required to undergo subsequent adjustments if they run afoul of this constraint. As we shall see in more detail in 2.3.2, English {eː, æ, ɔ, ɑ, aw, ay} normally become ɪ in Spaka (cf. (16)). When this process would create a violation of the Lax Vowel Constraint, though, {eː, ɑ, ay} instead become iː, and {ɔ, aw} become ɪw (Diehl and Kolodzey 1981: 415), neither of which outputs violates the constraint.

Parallel to this language game effect, there is ample evidence from repetition and attention studies that both children and adults are generally aware of the syllable phonotactics of their language. Brown and Hildum (1956), for instance, found that adults distinguish between legal and illegal syllables of English; preschoolers (Messer 1967) and even infants (Friederici and Wessels 1993; Jusczyk *et al.* 1993) can distinguish between phonologically legal and illegal syllables as well.

Closer investigation suggests, though, that not all subsyllabic sequences that are unattested in conventionally-occurring language data are assumed by speakers to be banned by the grammar. English is commonly believed to disallow syllable-initial sequences of ʃ + one or more consonants, for example (with the possible exception in Yiddishisms of ʃm (schmuck), ʃn (schnorrer), ʃl (schlep),[12] ʃw (schwing), ʃp (spiel), ʃt (schtick)), but given the stimulus in (9), a non-negligible number of English speakers who took our shm-reduplication survey (20 out of 356, or 6%) produced *Shmristmas presents*, with the seemingly illicit cluster ʃmr.

(9) Question 17 from Vaux and Nevins (2005)

> **Person A**: I never get enough Christmas presents. Nobody likes me!
> **Person B**: Christmas presents, _____! You should be happy with what you have.

The ʃmr cluster was even more popular with the simpler form *broom*: 32% of 414 respondents in a separate survey[13] favored *broom shmroom*, suggesting that (at least for these respondents) ʃmr is an accidental rather than systematic gap in the English lexicon.

Esper (1925), the first study I am aware of to use an artificial language learning methodology, takes the investigation of phonotactic knowledge one step further, asking whether such knowledge constrains the learning task. He asked test subjects to learn the names of 16 objects, each having one of four different shapes and one of four different colors (subjects were trained on 14 of the 16 possible

object-name associations and then tested on possibilities 15 and 16 to see if they had generalized what they had learned). Test subjects were separated into three groups, which were placed in three distinct experimental conditions. Group 1 was exposed to names such as naslig, sownlig, nasdeg, and sowndeg, where nas- and sown- coded color and -lig and -deg coded shape. Since these names each consisted of two morphemes that were phonotactically licit in English, the native language of the test subjects, this group could in principle simplify their task by learning eight morphemes rather than 14 names, plus the simple rule that the color morpheme preceded the shape morpheme in each name. Group 2 was also exposed to bimorphemic names, but unlike with Group 1, the morphemes were not phonologically legal for English, for example, -lgen and -zgub were shape morphemes. The names presented to the control, Group 3, contained no internal morphological structure, leaving subjects with no recourse but to learn 14 idiosyncratic names.

What Esper found after running his three groups was that, as expected, Group 1 learned their names much faster and more accurately than Group 3. Interestingly, the performance of Group 2 was similar to (and marginally worse than) that of Group 3, suggesting that learners' expectations about possible morpheme shapes based on their knowledge of English phonotactics were brought to bear (and in this case interfered) in the acquisition task. In fact, analysis of the errors of Group 2 revealed that they tried to make phonotactically legal morphemes from the ill-formed ones.

Esper's study suggests an important point: speakers have active knowledge of phonotactic restrictions on morpheme shapes, or Morpheme Structure Constraints.[14] His work therefore converges with the independent evidence for MSCs from Lac Simon discussed in Section 2.1.3.

Recent research has employed artificial language learning paradigms to demonstrate that both children and adults are able to acquire new phonotactic constraints, paralleling the finding by Dell *et al.* (2000) that test subjects conform to the phonotactics of toy grammars in their speech errors during attempts to produce forms in those languages. Onishi, Chambers, and Fisher (2002) asked adult test subjects to listen to a list of words produced by a toy grammar similar to that of English but possessing an additional syllabic restriction on the distribution of consonants. After familiarization with this list, subjects were presented with novel forms produced by the same grammar and asked to repeat them. The subjects were found to have taken significantly longer to repeat forms that violated the novel phonotactic constraint than forms that did not violate it, suggesting that they had successfully extracted and internalized the constraint. Chambers, Onishi, and Fisher (2003) found that infants showed the same effect when exposed to the same stimuli as the adults had been. Koo and Cole (2007) determined that test subjects can acquire not only first-order phonotactic generalizations (e.g. m is disallowed in syllable coda), but also second-order generalizations (e.g. coda obstruents are voiced after front vowels). Similar results with regard to acquisition of second-order phonotactic generalizations were obtained via a speech error elicitation paradigm by Warker and Dell (2006).

2.2.2 Movement, Reduplication, and Precedence Another topic of central concern to experimental investigators of ludlings and phonological theory is the encoding of linear precedence and its relationship to movement processes, reduplication, and truncation. A number of phonologists have noted resemblances between reduplication, truncation, and precedence-altering language games (Nevins and Vaux 2003; Idsardi and Raimy 2005; Yu 2004; Guimarães and Nevins 2006, among others). With the advent of OT certain researchers have proposed mechanisms to deal with ludlings such as Base-Game correspondence (Barlow 2001) or Base-Ludlingant Correspondence (Sanders 1999), implicitly appealing to the same mechanism that is claimed to drive reduplication. As mentioned in Section 2.1.2, ludlings appear to target a restricted set of anchor points. One finds similar restrictions on the anchor points targeted by iterative infixation games; for example, no ludling can require infixation before the third vowel, or before antepenultimate syllables, or perform iterative infixation on all syllables *except* the first.[15]

Iterative infixation, one of the most widespread types of language game, can be easily formalized using the Multiprecedence-and-Linearization framework of Raimy (2000). For example, iterative infixation that goes before every vowel, as shown in (8) for Ibenglibish,[16] can be represented in terms of a command to add precedence relations towards the sequence ajb from the preceder of each vowel and add precedence relations from the sequence ajb to each vowel. If such commands are limited to universal quantification, there is no way to have a ludling affect every vowel except the first (even though such a game might in principle be equally fun to play as others).

(10) Precedence relations for Ibenglibish (Idsardi and Raimy 2005)

$$\begin{array}{l} a \rightarrow j \rightarrow b \qquad\quad a \rightarrow j \rightarrow b \\ \nwarrow \quad \swarrow \qquad\quad \nwarrow \quad \swarrow \\ \#\rightarrow ɪ \rightarrow ŋ \rightarrow g \rightarrow l \rightarrow ɪ \rightarrow ʃ \rightarrow \% \end{array} \qquad \textit{English} > \textit{Ibenglibish}$$

Similarly, (11) shows a Multiprecedence-and-Linearization representation for the Brazilian Portuguese ludling Língua do Pê, for example, vela → vepelapa, where ɑ and ω denote abstract boundary symbols. In this game, a new precedence relation is added between every vowel and the consonant p, and between the consonant p and every vowel.

(11) Precedence relations for Língua do Pê (Guimarães and Nevins 2006)

$$\begin{array}{l} |\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega| \\ \qquad\quad \downarrow\uparrow \quad \downarrow\uparrow \\ \qquad\quad p_1 \quad p_2 \end{array}$$

Botne and Davis (2000) mention many other iterative infixation games, for example, the Spanish game Jerigonza (Piñeros 1998). Interestingly Botne and Davis argue for a third analysis in terms of imposition of a consonantal articulation on vowels, allowing them to account for the prevalence of labial consonants in iterative

infixation, which by their lights interferes the least with the vocalic articulation (though see Note 21 for problematization of this idea). The representation of iterative infixation in the precedence-based framework of Raimy (2000), combined with the notion of universal quantification, provides straightforward analyses of such patterns.

However, of more interest to the current discussion are "movement"-based languages games such as Pig Latin, which is often informally (and, for most varieties, incorrectly) characterized as "move the first consonant to the end of the word and suffix -ay." Such games can also be formulated in the precedence-based framework of Raimy (2000), though in somewhat more complicated terms. In actual fact, however, the "movement" involved in games like Pig Latin often leaves residue, failing to completely erase the original position of the moved elements. Nevins and Vaux's online survey of Pig Latin, for example (cf. Note 4), found that four out of 447 respondents transformed pig into pig-pay, with doubling of the p, while 13 changed tree into ree-tray, two into tree-tay, and one into tree-tray. (Interestingly, none transformed tree into tee-tray; I attribute this to the fact that no set of precedence instructions and valid anchor points can produce this sort of output.)

Similar facts can be found elsewhere, for example, in the syllable-swapping game Smoi represented in (12):

(12) Duplication in the Norwegian language game Smoi (Jahr 2003: 294)

jenta	→	tajent	'(the) girl'
banken	→	kenbank	'the bank'
flaska	→	skaflask	'the flask'

The fact that two instances of one or more consonants surface in such games, one in the base and one in the ludlingant, suggests that a purely movement-based account is incomplete or incorrect.

These kinds of results make sense in Steriade's (1988) model of reduplication, which involves full copy plus deletion. Extending her analysis to canonical Pig Latin, for example, the sequence <pig> + LUD[17] + -<ay> first undergoes full reduplication of the base to become pig+pig+ay, and then undergoes deletion of some sub-set of the segmental material in the base and the ludlingant. For most speakers of Pig Latin the set of segments deleted in the base and the ludlingant respectively is required to be complementary, resulting in the appearance of movement; for those who do not have this requirement, the copying underlying "movement" games is revealed. (Idsardi and Raimy provide an alternative account of the process that also avoids movement.)

Treiman and Danis (1988) found similar copying with children's syllabification games; for example, when asked to produce polysyllabic words backwards, children sometimes produced forms such as *lemon* → *mon-lem*. One could argue, though (as did Treiman and Danis), that this effect was a product of the intervocalic consonant being ambisyllabic.

Yu 2004, discussing Homeric Infixation (a language game made famous by forms such as *saxophone* → *saxomaphone* used in the TV show "The Simpsons"), argues that cases like *oboemaboe* involve compensatory reduplication. He proposes that *-ma-* must be placed after a trochee (e.g. *edu-ma-cation*) but cannot be final, and hence the *boe* of *oboe* must be compensatorily copied. Though promising, this analysis is silent about what speakers do with forms that contain more than one trochaic foot. Elfner and Kimper (2008) discuss examples with diddly-infixation that also involve copying, such as *wel-diddly-elcome* from *welcome*.

We have seen in this section that experimental methods and survey techniques can be successfully applied in tandem with natural and artificial language games to shed light on aspects of phonological competence and computation where we would otherwise be limited to natural language data that allow for incomplete and unsatisfactory inferences at best. In the case of phonotactics, for example, simple observation of distributional patterns in the naturally-occurring lexicon is not sufficient to demonstrate that speakers have active knowledge of them and employ it for phonological purposes; the convergence of ludling and acquisitional evidence on this point is therefore of essential importance. By the same token, extensive evidence from natural and artificial language games can clarify properties of phonological computation that are obfuscatorily underdetermined by natural language data, such as the fact that supposed movement processes actually involve reduplication.

2.3 Issues of Particular Relevance to Optimality Theory

Though most of the issues touched on thus far in this chapter are directly relevant to current work in phonological theory, for example with regard to the modeling of variation, absolute ungrammaticality, and acquisition, they were introduced for other purposes and should be of equal interest to phonologists of any stripe. In this final expository section of the chapter, I turn to language game data bearing on two areas of particular relevance to Optimality Theory: levels, rules, and rule ordering; and markedness and naturalness. We shall see that the ludling data in these arenas pose significant challenges for standard optimality-theoretic beliefs about the architecture and ontology of the grammar.

2.3.1 Levels, Rules, and Rule Ordering One of the most striking and potentially most important characteristics of language games is their ability to target different stages of the phonological derivation. This holds for (i) ludling processes in different languages, (ii) different ludlings within a single grammar, and (iii) a given ludling in the grammars of two different speakers.

An example of type (i) variation is Arabic -VVrb- infixation, which follows stress assignment in Meccan but precedes it in Hadrami (Walter 2003). Bagemihl (1988: 488) states that Tigrinya -gV- infixation precedes Tier Conflation, whereas Moroccan Arabic Permutation 2 follows it; similarly, Hebrew -goV- infixation precedes Pause Insertion, whereas Yoruba -gV- infixation follows it.

Type (ii) variation (cf. Bagemihl 1988: 503) occurs in Korean, where their syllable-reversing game mentioned earlier precedes the rules of fortition and resyllabification, but their iterative infixing game normally follows them (cf. (8)). Similarly, six of the 11 Salvador Brazilian Portuguese speakers tested by Guimarães and Nevins (2006) underapplied its Nasalization rule in (artificial) Língua do Ki but overapplied it in (natural) Língua do Pê, which can be interpreted in terms of Língua do Ki applying before (and hence bleeding) Nasalization and Língua do Pê applying after (and hence counter-bleeding) it.[18]

(13) Rule ordering in Brazilian Portuguese language games

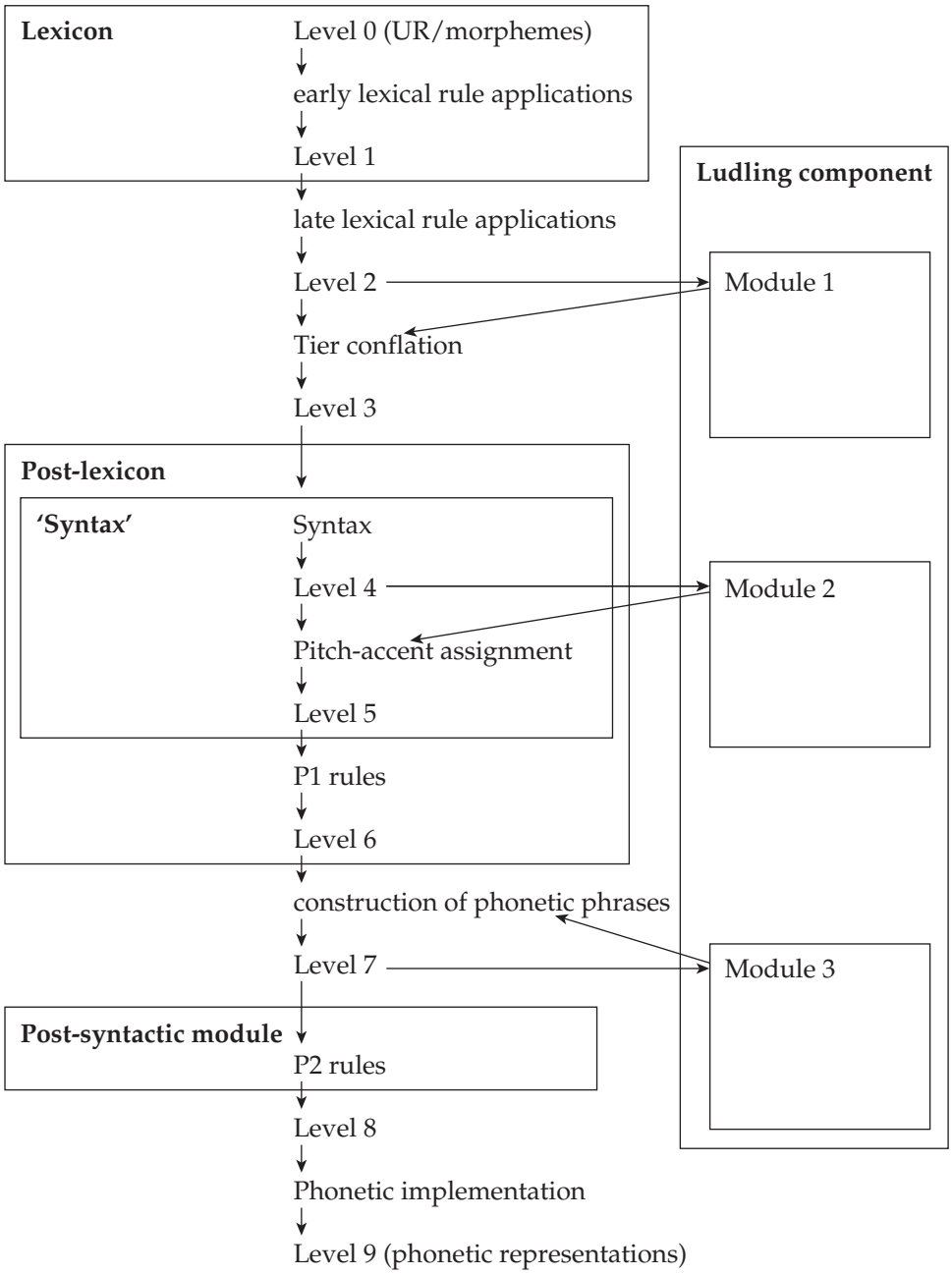
a.	normal Salvador BP output	[kãmiza]	'shirt'
b.	Língua do Ki output	[kãkimikizaki]	(underapplication)
c.	Língua do Pê output	[kãpãmipizapa]	(overapplication)
d.	compatible ordering:		
	i. Língua do Ki		
	ii. Nasalization		
	iii. Língua do Pê		

Type (iii) variation, where a given ludling is ordered differently in the grammar of two different speakers, is found most famously in the Cuna game Sorsik Sunmakke (Sherzer 1970; Bagemihl 1988: 508–511), but also in Kekchi Jerigonza (Campbell 1974; Bagemihl 1988: 511–513) and in Korean, where contrary to the normal order described earlier, some speakers apply the iterative infixing game before resyllabification. Similarly, three of the 11 speakers tested by Guimarães and Nevins (2006) ordered *both* Língua do Ki and Língua do Pê after Nasalization (p. 11), and 2 of the 11 speakers ordered both Língua do Ki and Língua do Pê before Nasalization (p. 12).

A fourth type of ordering variation, wherein a ludling optionally applies at two different points in the derivation within the grammar of a single speaker, is also conceivable but to my knowledge unattested.

So far in this section we have seen that, in derivational terms, ludlings can vary widely in terms of the stage at which they apply. Are there any limits on this variation? McCarthy (1991: 29) has suggested that ludlings apply at the end of the Word level and never precede lexical rules, but Bagemihl (1988) has shown that there is more variation than this in the point in the derivation at which ludlings can apply. Bagemihl (1988: 492) suggests (working within a Lexical Phonology framework) that ludlings can target outputs of three (and only three) stages of the phonological component as schematized in (14): Level 2 (i.e. immediately preceding Tier Conflation), Level 4 (post-lexical but structure-preserving, able to apply across word boundaries), and Level 7 (post-lexical and structure-changing, after phonological phrases have been assembled).

(14) Points in the phonological component at which ludling rules can apply, according to Bagemihl (1988)



If Bagemihl's generalization is correct, it works well in modular phonological frameworks such as Lexical Phonology (in both its rule-based and constraint-based manifestations, e.g. Kiparsky (1982a) and (2000) respectively, as well of course as Bagemihl's rule-based version in which the analysis in (14) is couched) and classic Rule-Based Phonology (Kenstowicz 1994a). It is not immediately clear, on the other hand, how such a generalization is to be captured in classic monostratal OT (Kager 1999) or even in the overtly derivational OT with Candidate Chains (McCarthy 2007).[19]

A related problem raised for OT by the behavior of language games is what Bagemihl (1988: 491) calls non-peripherality: "no ludling requires conversion earlier than Level 2 or later than Level 7. In other words, neither underlying nor surface representations are accessed by the ludling component." This appears at first blush to pose a problem for one of the central tenets of Classic OT, namely that phonological operations can only refer to URs and SRs, precisely the two representations to which ludlings, according to Bagemihl, do not refer. It remains to be determined, however, whether Bagemihl's generalization poses any actual empirical problems for Classic OT.

2.3.1.1 Spontaneous Opacity Perhaps a more serious problem for OT is the opacity effects that often surface in language games. The English-based sibling language Spaka, for example, possesses an exchange rule that transforms English ɪ to æ and vice versa (Diehl and Kolodzey 1981: 414). As we have already seen, -VVrb- insertion in Meccan Arabic renders its rule(s) of stress assignment opaque (Walter 2003), and Língua do Pê counter-bleeds Nasalization for some speakers of Salvador Brazilian Portuguese (Guimarães and Nevins 2006).

As pointed out by Vaux (2008), OT has at its disposal various devices that can *generate* the equivalent of opaque rule orderings, but existing algorithms for *acquiring* OT grammars provide no mechanisms for spontaneously generating opacity effects of this sort, and in fact predict that opaque configurations should not appear unless motivated by the primary linguistic data, which is not the case with spontaneous opacities of the sort catalogued here.

2.3.1.2 Local Ordering Another spontaneous and non-trivial phonological effect that emerges in language games involves Local Ordering. Anderson (1969) noticed that in Icelandic and Faroese certain rules appear to apply in different order for different forms, depending on which order is most transparent for that form. He dubbed this phenomenon Local Ordering; similar cases in other languages were identified shortly thereafter for Sundanese (Howard 1972) and Balto-Slavic (Darden 1978). A form of Local Ordering seems to surface in the Korean language game involving iterative infixation that was presented in Section 2.1.2.2, but in this case the ludling rule appears to be reordered not to maximize transparency, but rather to avoid sequences of three or more identical elements.

Recall that the game inserts after each vowel in a word a fixed consonant followed by a copy of the vowel, and deletes any Coda consonants following it, as in /k'aŋts^{h}oŋ/ 'hopping' → [k'apats^{h}opo]. The rules involved in the ludling

must follow the regular Korean processes of fortition and syllabification, as we saw in (7) and (8). Crucially for our purposes, the ordering of the ludling processes can change when the fixed consonant is identical to one of the consonants in the base form. If we insert -k'V- in /hakkyo/, for example, we expect to obtain *[hak'ak'yok'o], but what we actually get is [hak'akyok'o], where it appears that the ludling has bled fortition. Similarly, if we insert -mV- with a form such as /nim-i/, the ordering in (8) predicts *[ni.mi.mi.mi] but what we actually get is [ni.mi.i.mi], where the ludling appears to have bled word-level syllabification.

It seems, as I alluded to in Section 2.1.2.2, that something like an OCP constraint on sequences of identical syllables not otherwise motivated in the language is spontaneously surfacing in the language game. Whether this is a case of a hidden universal constraint surfacing or a local rule reordering being triggered by an OCP violation remains unclear, but we can be fairly confident of the basic fact that the language game has revealed a larger linguistic principle not present (or at least not visible) in the regular phonology of Korean.

What emerges from our discussion in this section is that ludlings present strong evidence for a wide range of phonological effects that are straightforwardly modeled in RBP but for Classic OT are at best problematic (as with Bagemihl's ordering generalization) and at worst unexplainable (as with variation in ludling order, spontaneous emergence of opaque rule ordering, and local reordering). Local reordering of the sort found in Korean is also potentially highly problematic for encapsulated derivational frameworks such as RBP, because it appears to require look-ahead power and trans-derivational comparison.

2.3.2 Naturalness, Complexity, and Markedness Another area of concern for OT where language games are particularly relevant involves meta-computational issues of naturalness. The rise of Optimality Theory has led to a resurgence of interest in naturalness and markedness, which are arguably captured more easily in OT than in RBP. McCarthy and Prince (1994) have observed, for instance, that markedness can be expressed directly in OT in terms of the number of violation marks for a given markedness constraint. Similarly, naturalness can be characterized in OT in terms of (depending on one's perspective) a configuration that can be generated by the universal constraint set CON, a configuration produced by a relatively high percentage of permutations of CON, or a configuration favored by a single constraint or a small set of interacting constraints.

But is phonology actually natural? And do unmarked segments necessarily emerge when all else is equal? Advocates of Optimality-Theoretic Dispersion Theory (e.g. Flemming 2004; Ní Chiosáin and Padgett 2009) say yes – one can predict the quality of phonemes in a system, including that of epenthetic segments, if one is given the quantity of members it contains. Much natural language evidence runs counter to these claims (see Disner 1983), and this evidence is complemented by the language game data in interesting ways. The children's secret language Spaka, for example, reduces the American English vowel inventory to the centralized system in (15) (Diehl and Kolodzey 1981: 414):

(15)

ɪ		i̵	i̵y
		ə	
æ		ʌ	
		ay	

As Diehl and Kolodzey point out, a seven-vowel system clustered around the center of the vowel space in this way is hardly what we expect if vowel systems are optimally dispersed, unless ease of articulation trumps maximization of contrast (which, however, will still not yield the system in (15)).

The choice of epenthetic segments is even more problematic, in both natural languages and language games. The latter appear to draw freely from the inventory of segments in the language on which the game is parasitic, such as /n/ and /m/ in Spaka (Diehl and Kolodzey 1981: 416, 417), /f/ and /g/ in Ẹna (Iṣọla 1982), /p/ in Língua do Pê (Guimarães and Nevins 2006), /h/, /s/, or /ʒ/ in Moroccan Arabic (Heath 1987), /ʌb/ in Ubbi Dubbi (Ribeiro 2000).[20] In iterative infixation games such as Ubbi Dubbi and Língua do Pê, the fixed melodic material can be argued to be a separate morpheme and thus not generated by the phonology. The same does not hold, though, for Spaka m- and n-insertion, which are purely phonological (see Diehl and Kolodzey 1981 for discussion).

The unnaturalness of most ludling processes also poses difficulties for feature-based natural theories of phonology such as Natural Phonology, OT, and the highly constrained feature geometry-driven RBP of McCarthy (1988). Theories of this type are specifically designed to rule out unnatural processes, on the reasoning that such processes do not exist in natural languages; their existence in language games is therefore a bit of an embarrassment. How is one to account for the Spaka vowel neutralization rules in (16), for example?

(16) Vowel neutralization rules in Spaka (Diehl and Kolodzey 1981: 414)

a. eː, æ, ɔ, ɑ, aw, ay → ɪ
b. ɪ, ʊ → æ
c. ɛ, ʌ → i̵
d. uː, oː → ʌ

RBP of SPE vintage deals with such rules fairly well (though the fact that some of the rules do not submit happily to feature-based formulation is of some concern); theories which strive to limit their predictive power to what we find in well-known natural languages do not.

Thus far natural languages and ludlings converge on the same facts about phonological systems: they are not always natural. Recent work on the acquisition of artificial language games has enabled us to move a step beyond what we can see reliably in natural languages, however. In a laboratory setting it is possible to independently manipulate naturalness, complexity, and frequency, variables which tend to co-vary in naturally-occurring phonological systems. These variables were first manipulated in an experimental setting by Pycha *et al.* (2003), who investigated the acquisition of artificial grammars containing harmony rules whose

characteristics varied in complexity in one condition, and in naturalness in a second condition. They found that test subjects learned the simpler harmony rule (which manipulated a single phonological feature) more successfully than a more complex one (which required manipulating a larger number of features), but they did not fare better with the more natural rule than the less natural one. Similarly, Wilson (2003) found a learning bias for processes of consonant harmony and disharmony relative to arbitrary (i.e. more formally complex) suffix selection processes. Seidl and Buckley (2005) found that common, phonetically grounded patterns (such as intervocalic voicing of stop consonants) were not learned significantly more successfully than rare, phonetically arbitrary patterns (such as intervocalic stop devoicing). Moreton (2008) tested adults' learning of cross-linguistically favored height assimilation of vowels (in which vowels in adjacent syllables agree in height) with an unattested, distant voice assimilation pattern in which the initial consonants of adjacent syllables agree in voicing, and failed to find a bias for the former, suggesting that relative cross-linguistic frequency does not correlate with ease of acquisition. Koo and Cole (2006) obtained similar results for the acquisition of non-local phonotactic constraints with regard to frequency and phonetic naturalness. Pycha, Shin, and Shosted (2006) studied the acquisition of three types of consonant assimilation rule, one natural (regressive), one unnatural (progressive), and one highly complex (variable in direction). They found no significant difference the learnability of the natural versus unnatural patterns, but a significant difference in between the formally simple and complex patterns. Similar results were obtained by Skoruppa and Peperkamp (2008), who found that segmental alternations differing in only a single feature were acquired more successfully than alternations differing in three features.

What the phenomena discussed in this section show is that language games can shed light on the phonological component qua computational system, stripped of the historical strata that (according to Ohala 2005 among others) produce a patina of naturalness. Once the products of history have been removed, phonological learning is revealed to operate in terms of formal complexity defined in terms of featural representations, rather than on the basis of frequency or naturalness (in a non-featural, phonetic sense).

3 Summary

In this chapter I have attempted to complement existing typological surveys of language games (most notably Laycock 1972 and Bagemihl 1988) and conventional analyses of their workings by updating and expanding our purview to include artificial language learning experiments and ludling-like processes such as English shm-reduplication and Expletive Infixation, and by considering what these reveal about the architecture of the phonological component, the nature of phonological learning, and the implications of these for general linguistic theory. I have suggested that the range of ludling phenomena surveyed here reveals a number of recurrent and significant properties:

(i) *Variation.* Learners vary widely in the analyses of underdetermined ludling data that they entertain, suggesting that the learning algorithm must have a stochastic component rather than being entirely deterministic. A non-trivial number of the hypotheses considered by learners are "crazy," suggesting that phonology is not entirely natural in any conventional understanding of that term. An individual hypothesis may also allow for significant variation in the range of inputs it targets and outputs it produces (hypervariation), suggesting that the phonological component is able to make use of existential operators in an open-ended way that may be difficult to incorporate in OT.

(ii) *Constraint.* Though the range of attested variation is much larger than predicted by most current theories, it is far from being as extensive as one would expect if all hypotheses compatible with a given data set were considered. Learning appears to be limited to elements and operations provided by UG such as anchor points and sonority-driven syllabification, and constrained by a set of priors including knowledge of language-specific Morpheme Structure Constraints, possible feature interactions (cf. Wilson (2006) on palatalization on vowel height and Becker *et al.* on consonant voicing and vowel backness and height), and OCP-type principles (as surfaced in shm-reduplication and Korean iterative infixation).

(iii) *Abduction*. Learners often generalize beyond what is minimally required by the primary linguistic data, taking free rides on alternations observed elsewhere, extending patterns to new segment classes, and extracting first- and second-order phonotactic generalizations at both underlying and surface levels from non-alternating distributional patterns. This aggressive learning can be quite rapid as well, as one can see anecdotally in the postulation of Pig Latin rules from the stimulus *ig-pay atin-lay* alone, and experimentally from the Goldilocks Rule identified by Gerken and Bollt 2008 (cf. also Fast Mapping in the acquisition of lexical semantics (Carey and Bartlett 1978)).

Adjusting our phonological theory to take account of these phenomena revealed by ludlings should enhance our understanding of naturally-occurring phonological phenomena as well.

ACKNOWLEDGMENTS

Thanks to Greg Iverson, Neil Myler, Claire Nance, Andrew Nevins, Eric Raimy, Cilene Rodrigues, and Andrew Wolfe for help with drafts of this chapter.

NOTES

1 Note that by adding to traditional language games findings from invented games for existing languages, toy grammars, and artificial language learning experiments, I cast

the net wider than Bagemihl (1995), who defines a ludling as "a language which meets the following criteria: (1) its morphological system is limited to one or more operations drawn from the following (a) infixing/affixing, (b) templatic, (c) reversal, (d) replacement; (2) its affixes (whether fully specified or defined only in prosodic or melodic terms) are limited to one or at most a handful of lexical items; and (3) its morphology is semantically empty." I believe that the wider scope is justified by the fact that all objects of our investigation involve manipulation of phonological elements that move beyond what can be observed in conventional phonological systems. I am also unconvinced by Bagemihl's assertion that ludlings have no morphological or semantic content, but this question is not relevant for the phonological purposes of this chapter.

2 Zwicky (1981: 598) identifies as external evidence in phonology anything other than "data on the cooccurrence and alternation of linguistic elements in some language, as well as such systematic considerations as formal simplicity, economy, and the like."

3 For recent formal analyses of individual language games see, for example, Davis and Hammond (1995), Itô, Kitagawa, and Mester (1996), Piñeros (1998), Sanders (1999), Botne and Davis (2000), Kaun and Harrison (2001), Barlow (2001), Nevins and Vaux (2003), Moreton, Feng, and Smith (2005), Rizzolo (2007), Elfner and Kimper (2008), Yu (2008).

4 http://php-dev.imt.uwm.edu/prjs/markj/projects/fll_surveys/piglatin/, accessed August 5, 2009.

5 Available at http://www.tekstlab.uio.no/cambridge_survey/views/1099.

6 Participants (whose number is unspecified by Nevins and Endress) were informed that they would witness a Martian rite in which a chief Martian speaks a word and a subordinate Martian must then reply appropriately. Participants were presented with 25 trials, in which the chief Martian uttered a trisyllabic sequence and the subordinate Martian replied with the same syllables but in reverse order. After familiarization, participants were informed that they would witness the rite, now with the same chief Martian and another Martian who has not learned the rite as well. Subjects were told to judge on a scale from 1 to 9 whether the new subordinate Martian's response conformed to the rules of the rite. They then completed 20 trials in which the chief Martian uttered a four-syllable sequence, and the new subordinate Martian replied with the same syllables in one of four different orders.

7 For instance, Raimy derives overapplication of the Malay rule that nasalizes post-nasal vocoids in reduplicated forms such as /aŋen + RED/ 'unconfirmed news' → [ãŋẽn ãŋẽn] 'germs' (not expected *[aŋẽn ãŋẽn]) via a reduplication rule

$$\# \rightarrow X \ldots X \rightarrow \%$$

operating on representations such as #→a→ŋ→e→n→% and the Uniformity Parameter set to Off.

8 Here, as throughout the chapter, I consider phenomena such as English Expletive Infixation, Homeric Infixation, and *shm*-reduplication and Armenian and Turkish intensive reduplication and m-reduplication to be language games for phonological purposes, though their pragmatics may differ somewhat. For more on the idea that marginal phonological processes of this type can be treated as language games, see Zonneveld (1984) on English Expletive Infixation and Bagemihl (1988: 498) on Tagalog *um*-infixation.

9 Moon (1998) notes that some speakers order the ludling before syllabification and some after, but all speakers order it after fortition.

10 Specifically, they were more willing to insert the infix inside clusters of rising sonority than clusters of falling sonority (see Pierrehumbert and Nair 1995: 101, Table 2).
11 McCarthy (2005) postulates a mechanism that allows for free rides in OT, but it is not clear that this mechanism is compatible with Lexicon Optimization.
12 Fans of Celtic sports may also know the loanword sliotar 'hurling ball' [ʃlɪtə(ɹ)].
13 http://survey.net.nz/results.php?7b71e88ecfa1f070b7199fd4f5fbf785, accessed August 4, 2009.
14 Esper's study effectively studies stem morphotactics. Applying his findings to morpheme acquisition runs afoul of the fact that many languages possess bound morphemes whose phonological content in isolation is phonotactically illegal, for example, the Quechua past tense morpheme *-rqa* (Neil Myler, personal communication).
15 It is interesting that there are cases where iterative affixation targets all syllables except the final (Yu 2007a: 2008), which is easily treatable via an Alignment constraint in OT but may require more subtle machinery in Raimy's model.
16 Also known as Goose Latin (Pollack, Gandour, and Sorensen 1987).
17 LUD is Sanders' (1999) term for a ludlingant morpheme, analogous to RED for a reduplicant morpheme.
18 Cf. Bagemihl (1988: 505) for a similar case of regular phonological rules sandwiched by ludling rules in Mandarin May-ka language. For extensive discussion of the exact environments in which Nasalization applies in Salvador, and justification of vowel nasalization being derived rather than underlying, see Guimarães and Nevins (2006).
19 It might be worth adding here that OT is generally incapable of accounting for several of the core properties of lexical rules (or their converse in post-lexical rules) noted in classic work on Lexical Phonology, such as structure preservation, cyclicity, exceptionality, and conscious awareness.
20 Botne and Davis (2000) claim that iterative infixation games tend to employ labial consonants and attribute this to the fact that a labial articulation does not disrupt the (presumably coronal or dorsal) vocalic gesture in the middle of which it has been inserted, but this analysis has difficulty accounting for the abundance of non-labial (and indeed coronal and dorsal) consonants inserted in such games.

23 Loanword Adaptation: From Lessons Learned to Findings

CAROLE PARADIS AND
DARLENE LACHARITÉ

1 Introduction

Why are French *élan* [elɑ̃] and *nuance* [nyɑ̃r] pronounced [elæn] and [nuans] in English? Why are English *credit* [kɹɪɛdɪt] and *desk* [dɛsk] pronounced [kɯredʒitto] and [desɯkɯ] in Japanese? The answer is that these loanwords/borrowings contain foreign sounds or phonological structures (referred to as malformations) that the borrowing languages (L1) do not permit, that is, do not have in their own phonological inventories. These malformations, if they are not tolerated (imported), have to be adapted, that is, made to conform to the sounds and structures of the L1. For example, English does not have phonemic nasal vowels or front rounded vowels, so French [ɑ̃] and [y] are replaced by English native [æn] and [u]. As for Japanese, it has neither branching onsets, nor codas (at least not those requiring a place node, cf. Itô 1986), so the English examples above undergo vowel insertion in several places. Whether these replacements, which are part of the process of loanword adaptation, are based on phonology or phonetics is still debated. We will return to this issue shortly, but state at the outset that we share the view that loanwords have a lot to teach phonologists about the functioning of phonological constraints. Some of the reasons for our phonological bias will be presented in this chapter.

The interest of phonologists in loanword adaptation has mushroomed since the emergence of constraint-based theories, because the patterns of sound changes in loanwords offer abundant positive evidence for the phonological constraints that a language places on output. What do loanwords provide for phonology that native words do not or cannot? Even where native processes furnish evidence for a particular constraint in a language, there are often few pertinent items on which

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu
© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

to base generalizations. Borrowing – whether we like it or not – is normal language contact behavior and most languages have a significant number of borrowings. Thus, for almost any language, it is relatively easy to assemble a corpus large enough for adaptation patterns (by extension a variety of constraints) to emerge. Also, foreign lexical items are unencumbered, or far less so, by morphological residue than native words. The more evidence there is for a given constraint, in terms of the number of items to which the constraint applies and in terms of the morphology-free nature of the constraint's effects, the more confident phonologists can be about the shape of that phonological constraint. A language's reactions to loanword input containing foreign sounds or structures reveal a lot about the phonology of that language. In a broader perspective, by studying the loanword adaptation patterns in many different languages, phonologists can get a better idea of how phonology functions cross-linguistically and universally. This is why loanword adaptations are so important: they open a larger window onto the range and functioning of phonological constraints. It is true that when phonologists study established loanwords, as opposed to elicited forms or nonces, they are not working directly with online processes. Nonetheless, loanwords, especially recent ones, still provide some of the best evidence for online phonological processes, because the results of adaptation allow phonologists to easily deduce the online processes, while avoiding the inconveniences of elicited data (which are addressed later in the chapter).

This chapter presents an overview of what the intensive and extensive study of loanword adaptation reveals. The data and statistics discussed here come predominantly from Project CoPho's loanword database, which includes over 12,000 borrowings, mainly from English and French, in several different languages.[1] Because the methodology used and the numerical weight of a database have a profound effect on the conclusions one can legitimately draw about loanword adaptation, this chapter devotes a good deal of attention to discussing the gathering, classification and analysis of the data and the compilation of the statistics.

2 Overview of the Main Lessons Learned

One of the first lessons that an extensive study of loanwords teaches is the necessity to consider a corpus containing several hundred borrowings and to compute precise statistics for adaptations, because particular adaptations do not reveal themselves as exceptions or rules until there is a broad statistical basis for comparison. Exceptional cases tend to draw one's attention, and familiar patterns to fade into the background of awareness. These tendencies can leave the impression that unusual changes in loanwords are more common than is actually the case. For instance, in CoPho's corpus of English loanwords in Japanese, it initially seemed that English /kæ/ and /gæ/ sequences were nearly always adapted as /kja/ and /gja/, an impression shared by Dohlus (2005).[2] However, the study of the adaptation of a large number of English loanwords containing

/kæ/ and /gæ/ shows that adaptation as /kja/ and /gja/ represents a situation where a less usual adaptation pattern commands undue attention. When a good number of pertinent malformation cases are considered and the statistics drawn, one finds that this is a minority pattern. In the CoPho database of English loanwords in Japanese, adaptation to /ka/ and /ga/ occurs in 55/79 cases (69.6%). In only 24/79 cases (30.4%) does adaptation involve glide insertion. Thus only by having a sufficient number of cases can one begin to see patterns in the data and to distinguish the usual adaptations from those that are less common, or unusual.

Another important lesson that the study of loanwords teaches is that to reach valid conclusions about how sounds and structures are adapted by the borrowers of a given language, one needs to be sure that a given form entered directly from the source to the borrowing language. Indirect borrowings, that is, those L2 items that entered the L1 via another language, must be avoided, since borrowings that come through this channel risk confusing the adaptation strategies used by the L1 with those used by the third language. Needing to be sure of provenance heavily favors working with first-hand data and encourages wariness of secondary sources, unless and until their reliability can be established. For example, Schütz (1978) contains a good-sized corpus of English loanwords (about 800) in Fijian, making it tantalizing to use for contemporary phonological analysis of loanword adaptation. However, Schütz (1978: 10) says explicitly that English-speaking missionaries imposed many of the adaptations found in the loanwords provided. In other words, the adaptations reflected in the data are, in many cases, the missionaries' guesses as to how the word would best fit into Fijian, a language they knew only as non-native speakers, sometimes with a low level of proficiency. Unfortunately, Schütz does not indicate which loanwords have been adapted by the missionaries, rather than the Fijians themselves. Thus, this corpus would be an unreliable source on which to base generalizations about the Fijian adaptation of English loanwords.

In analyzing loanword data, it is also vitally important to distinguish between loanword adaptation, per se, and native processes. To do this, one must have sufficient reliable information about both the source and borrowing languages at the time the words were borrowed. The study of English loanwords in Old Quebec French (OQF), the French spoken in the province of Quebec in the nineteenth and early twentieth centuries, illustrates this (see Paradis and LaCharité 2008 for an in-depth analysis of loanword adaptation in OQF). At first glance, the corpus seems to contain a great number of unusual adaptations and forms. For example, English /n/ sometimes surfaces as /l/ in OQF (Eng. *tenement* [tɛ́n̲əmənt] → OQF [tɛl̲əmɛn]); English /l/ as /j/ (Eng. *overalls* [óvəɹɑ̀l̲z] → OQF [ãvʀaj]. However, extensive research into the phonologies of OQF and the English of the period revealed that most so-called aberrant forms and adaptations were best explained as stemming from native French or English processes of the epoch. They did not result from loanword adaptation, per se. For instance, there was a common *l-n* alternation in OQF (that still occurs in a few words in contemporary Quebec French; for example, *caleçon* /kal̲sõ/ 'briefs' and *à l'envers* /al̲ãvɛʀ/ 'upside

down' are still pronounced [kan̲sõ] and [an̲ɑ̃vɛʀ] by several Quebec speakers) and word-final /l/ was often realized as [j] in OQF (*métal* /metal̲/ 'metal' → OQF [metaj̲]). The apparent unpacking of the front round vowel /y/ that corresponds to /ju/ in French loanwords in English (e.g. French *butte* [by̲t] → English [bju̲t]) also illustrates the need for reliable phonological information about the source and borrowing languages and the need to distinguish between loanword adaptation and native processes. One might think that the properties of the French front round vowel, which is a malformed sound from the point of view of contemporary English, are perceived as two distinct sounds, a front one and round one, since the combination in a single sound is prohibited in the L1. In other words, unpacking might be attributed to reliance on phonetic, as opposed to phonological information about the malformed sound. However, this viewpoint is sustainable only if unpacking is a process of loanword adaptation. Yet there are several arguments suggesting that English /ju/ sequences developed from a native diphthongization process of Middle English, rather than resulting from loanword adaptation (Paradis and LaCharité 1997; Paradis and Prunet 2000). For one thing, Middle English still had front round vowels, so these would not have needed to undergo adaptation on the grounds of being foreign sounds. Another reason for thinking of this as a native process is that glide insertion affected native English words containing /u/, such as *few* [fju] and *knew* [nju], not only French borrowings. Also affected were loanwords that never contained a front round vowel (e.g. French *coupon* [kupõ] came to be pronounced [kju̲pɑn], a pronunciation that, although lacking prestige, still exists in some dialects of English. In contemporary *non-djuwty* dialects of English, including standard North American English, yod insertion occurs before /u/ if the preceding consonant is a non-coronal as in *cute* [kʲute], though there are many exceptions (compare *cute* [kjut] with *coot* [kut]). English now commonly inserts /j/ before /u/ in words, including borrowings, that never contained a front round vowel, for example, Spanish *Cuba* /kuba/ yields English [kju̲bə]. In sum, the appearance of the glide in borrowings is more reasonably attributed to a native process that inserts /j/ before /u/, which is English's overwhelmingly dominant adaptation of French /y/.

As the above example illustrates, it is advantageous to know as much as possible about diverse aspects of the borrowing and source languages, and the social and historical context in which the borrowing occurs, even if phonology is the central interest. In language-contact phenomena, which include borrowings, the possibility of non-linguistic factors interfering with phonology must be considered. Borrowing can be an emotionally charged social issue, and social constraints and pressures may have an impact at the phonological level. In some situations, adapting all foreign sounds and structures may be viewed as insulting; in others it might be considered obligatory (Grosjean 1982). Also, it must be borne in mind that the study of borrowing is necessarily multidisciplinary; one has to consider pertinent findings in L2 acquisition, sociolinguistics, and history, to name just a few of the areas that borrowing routinely intersects.

If the factors mentioned in this section, for example, native phonological processes, indirect borrowing, passage of time, and the sociolinguistics of borrowing, are not taken into account and if we do not work with sufficiently large corpora and precise statistics, we might be left with the impression that phonetic approximation is much more active in loanword adaptation than is really the case. However, there is an even more fundamental reason that phonetic approximation is so readily invoked. It is widely believed by linguists as well as non-linguists that the main port of entry of loanwords into a language is the monolingual, who – confronted by a foreign concept – is compelled to pick up the foreign word for it and must do his or her best to perceive and reproduce the foreign word. This would result, of course, in phonetic approximation, since the borrower knows nothing of the L2 phonology. This brings us to the major debates in loanword adaptation.

3 Two Contemporary Major Debates in Loanword Adaptation Studies

There are currently two major debates in loanword adaptation that we will briefly address. The first, as just mentioned, is over the question of whether loanwords are adapted mainly via phonetic approximation or phonology. The second concerns the influence of orthography in loanword adaptation.

With respect to the first debate, we all agree that a foreign sound is adapted to the closest L1 sound. The question is this: Is "closest" determined on phonetic or phonological grounds? LaCharité and Paradis (2005) show that these are distinct, yielding different outcomes in many cases. In the phonetic approximation view, loanwords are adapted by monolinguals, or bilinguals operating in monolingual L1 mode, who thus do not have access to L2 categories (phonemes) and structures. They are therefore forced to adapt phonetic L2 outputs and map them directly onto L1 categories and structures without taking L2 categories into account. This is the essence of phonetic approximation (see Silverman 1992; Yip 1993; Kenstowicz 2007; Peperkamp and Dupoux 2002, 2003; Kang 2003; and so on, for in-depth discussions of and analyses from the phonetic approximation perspective).

The alternative view is that loanwords are adapted by bilinguals who are operating in bilingual mode. They thus have access to L2 categories (phonemes) and structures. The L2 phonetic output is used by bilinguals only to access L2 categories and structures, in order to transfer them into those of L1.

The different predictions of the two views can be illustrated by the case of English voiced stops in French. The Voice Onset Time (VOT) of English voiced stops overlaps with the VOT associated with voiceless stops in French, so English /b/ corresponds phonetically to French /p/. The phonetic approximation stance predicts that English /b/ will be adapted to French /p/ in large proportion, whereas the phonological stance predicts that English /b/ will be systematically

categorized as /b/, in spite of differences in VOT. An example of possible phonetic approximation is the adaptation of English *blackball* /blækbɑl/ as OQF [plakbɔl][3] instead of the phonological adaptation [blakbɔl]. Another example of the different predictions of phonetic approximation versus categorical adaptation is provided by the case of English /r/ in Japanese. The English rhotic /r/ ([ɹ]) is perceptually closer to Japanese /w/ than to the Japanese rhotic, which is realized as a flap (Mochizuki 1981; Best and Strange 1992; Yamada and Tohkura 1992; Guion *et al.* 2000). Phonetic approximation predicts that English /r/ will be adapted as /w/ in Japanese, whereas the phonological stance predicts that English /r/ will be identified as a rhotic and consequently replaced by the Japanese rhotic in English loanwords in Japanese, despite great phonetic differences between the two rhotics.

One of the reasons that we support the phonological stance is that phoneme confusions such as those just illustrated are either non-existent or extremely rare in the CoPho loanword database. In other words, English voiced stops /b, d, g/ are interpreted as /b, d, g/ in French, Italian, Spanish, Portuguese, and so on, despite important VOT differences. The same is true of the English rhotic /r/ ([ɹ]), which is interpreted as a rhotic (as opposed to the glide /w/) in Japanese, French, Spanish, Italian, and so on, languages whose own rhotics are phonetically very different from that of English (compare English [ɹ] with French [ʁ], [ʁ] or [r]).

Although this chapter will not focus on developing all the arguments in favor of the phonological view (for this we refer the reader to, for example, Ulrich 1997; Paradis and LaCharité 1997, 2001; Jacobs and Gussenhoven 2000; Davis 2004; LaCharité and Paradis 2005), we would like to mention that the phonological stance is supported by external evidence. For example, by showing that bilinguals are the online borrowers and adapters (e.g. Poplack *et al.* 1988; Field 2002), sociolinguistic studies dispel a common myth of loanword adaptation whereby the monolingual is confronted by a foreign concept and is compelled to do his or her best to perceive the foreign word for it. The phonological view is much more straightforwardly compatible with the finding that borrowers are generally bilingual. The phonological view is also consistent with studies showing that becoming bilingual has a profound effect on the perceptual system, such that, with increased exposure to the L2, speech sound perception becomes phonological and more like that of native speakers of the L2 (e.g. Williams 1979; Flege and Eefting 1987; Nathan 1987; Best and Strange 1992; Escudero 2005). Numerous studies also show that bilinguals or multilinguals cannot entirely disengage their known languages (Dewaele 2001; Dijkstra 2003; Jessner 2003, 2006; etc.), so there is no such thing as monolingual mode unless one is monolingual. In short, our bias is phonological, not just because that view best accounts for what we see within the Project CoPho loanword database, but also because it seems most compatible with research in related disciplines that study borrowing and other language-contact phenomena. Although we will not concentrate on the phonetics versus phonology debate, it will be referred to throughout the chapter, in the

course of presenting our findings and analyses. Where relevant, additional justification will be offered for the phonological stance.

Another debate in the loanword literature concerns the extent to which loanword adaptation is based on orthography. Despite the fact that this debate has received less formal treatment than the phonetics-phonology one (but see, e.g. Lovins 1975; Shinohara 1997; Goulet 2000; Lamoureux 2000; Paradis and Prunet 2000; Vendelin and Peperkamp 2006 for discussions of orthography influence), it is quite widely believed that spelling influences loanword adaptation to a notable extent. Orthography influence is at work when the sound adaptations of foreign words are referenced to the usual sound-spelling correspondences of either the source or the borrowing language, rather than being based directly on either the phonetic or phonological properties of the foreign word. For example, English *building* /bɪ̱ldɪŋ/ and *cutter* /kətəɹ/are often pronounced as [bɣ̱ldiŋ] and [kytəʀ] rather than [bildiŋ] and [kətəʀ] in French, most likely because the grapheme <u> in French represents the front round vowel /y/. However, we will show in Section 5.2 that, in the CoPho loanword database, orthography influence is statistically marginal.

4 Description of the Project CoPho Loanword Database and Methodology

Since most of the conclusions presented here are based on the study of the corpora in the Project CoPho loanword database, we begin by introducing those corpora and providing general statistics. The general corpora are presented in Section 4.1 and the targeted corpora, which are smaller corpora compiled to address specific questions and hypotheses, are presented in Section 4.2, along with an overview of the Project CoPho's directory of segmental changes. In presenting Project CoPho's corpora and tools, we define our terminology and elucidate what we believe to be important aspects of our methodology.

4.1 General Corpora

The bedrock of our views and conclusions about the treatment of foreign sounds and structures is the study of CoPho's general corpora, over 12,000 loanwords in 13 corpora of loanwords from English (in Old Quebec French, Parisian, Montreal and Quebec City French as well as in Mexican Spanish, Japanese, and Calabrese Italian) and from French (in Canadian English, Moroccan Arabic, Kinyarwanda, Lingala and Fula). General numerical information about these corpora is given in (1).

(1) Project CoPho's general statistics including segmental, syllabic, metrical, and stress malformations from thirteen corpora of French and English loanwords (July 29, 2009)

Corpus	*Borrowings*	*Forms*	*Malformations*
English borrowings in			
Old Quebec French	485	597	489
Parisian French	901	2,576	3,153
Quebec City French	949	2,416	2,434
Montreal French	949	2,248	2,285
Mexican Spanish	1,045	1,514	3,137
Mexican Spanish II	1,034	2,342	5,645
Japanese	1,167	2,991	7,760
Calabrese Italian	2,161	5,191	14,740
French borrowings in			
Canadian English	674	1,667	1,034
Moroccan Arabic	1,127	2,685	4,275
Kinyarwanda	756	2,130	4,639
Lingala	672	1,917	3,734
Fula	532	1,081	1,118
All corpora	12,452	29,355	54,443

Borrowings were gathered from documents and spontaneous speech, and, except for the corpus of OQF, their pronunciations were verified with a minimum of three L1 speakers. Each pronunciation of a loanword provided by an L1 consultant constitutes a "form." For instance, the corpus of English loanwords in Japanese consists of 1,167 loanwords whose pronunciations were verified with three native speakers of Japanese. This produced 2,991 forms. The fact that there are not quite three times as many forms as there are borrowings follows from the fact that not all consultants know every borrowing. The number of malformations (recall that a malformation is a foreign sound or structure that is not permitted in the L1 core grammar) is based on forms. In the case of Japanese, the 2,991 forms contain a total of 7,760 malformations. The great number of malformations relative to the number of forms is accounted for by the fact that a single borrowing often contains more than one malformation.

With the exception of the corpus of English loans in Old Quebec French, each of the general corpora contains not less than 1,000 malformation cases. The average for each CoPho corpus is 3,000 to 4,000 malformations, but some corpora contain up to 5,000, 7,000 and even 14,000 malformations. In sum, the conclusions

about loanword adaptation presented here are based on the study of a total of 54,443 malformations (32,297 segmental, 16,324 syllabic and metrical, and 5,822 accentual) from 29,355 loanword forms contained in Project CoPho's database.

The borrowings of CoPho's database are established loanwords that are no older than 200 years.[4] In fact, it is primarily some French loanwords of the Canadian English corpus that are that old. The English loanwords of the Old Quebec French (OQF) corpus are also relatively old since OQF refers to the language spoken in Quebec in the nineteenth and early twentieth centuries (though these borrowings were not old at the beginning of the century, when they were collected and pronounced by the more than two hundred informants consulted by Rivard and Geoffrion, and whose pronunciations are reported in the *Glossaire du parler français au Canada* published in 1931). Apart from these two corpora, the vast majority of the loanwords in the database were borrowed after World War II, and many are much more recent.

As just noted above, we verified their pronunciations with a minimum of three L1 speakers, none of whom (nor their parents) was fluent in L2. All the pronunciations were tape-recorded. Subject to considerations to be discussed in Section 4.2, we tried to select consultants from different backgrounds and age ranges, to minimize the number of borrowings unknown to any of the speakers. Note that our study is not sociolinguistically but phonologically oriented. This means that our purpose is not to draw as faithful a picture as possible of borrowing's uses and conventions in a given community, but to understand how the mental categorization system deals with phonological malformations cross-linguistically. We thus intentionally favor a small number of consultants (three to four) for each corpus and large corpora of borrowings from many different languages, instead of a large group of consultants for one or a very limited number of languages.

Distinguishing forms from loanwords is important, not just because a given consultant does not necessarily know every single loanword, but also because different consultants, or even a single consultant, may pronounce a given loanword in more than one way. Thus, by focusing on forms, our analyses take into account all the variant pronunciations of our consultants and disregard any loanwords whose pronunciation we are unable to authenticate. Forms were elicited mainly through tasks such as definitions, paraphrases, fill-in-the-blank phrases, pictures or drawings. The loanwords were not pronounced by the interviewers and, to minimize the possible influence of orthography, the words were not presented in written form either. On occasion, however, to accelerate a consultant's lexical access to borrowings referring to notions that were abstract or especially difficult to paraphrase, such borrowings were shown to the consultants, in groups of 20, before the interviewer applied the normal elicitation strategies just mentioned. The consultants' pronunciations were tape-recorded and subsequently transcribed in IPA notation. The transcriptions were independently verified by at least one other phonologist or phonetician before being entered in a FileMaker database. Malformations within each form were identified, analyzed, classified, and statistically compiled. Main phonological patterns of adaptation were established and divergent ones carefully examined to detect non-phonological influences such as

orthography, analogy (false or real, partial or complete), and phonetic approximation. If an adaptation was straightforwardly handled by phonological constraints and repairs, we considered it phonological, even if it was also compatible with a perceptual view because, as mentioned previously, our bias is phonological.

Sound adaptations that were identified as non-phonological were compiled apart from those that are phonological. Note that divergent adaptation patterns are not necessarily non-phonological. Even though there is, in the vast majority of cases, only one single phonological adaptation pattern for a given malformation within a language, it sometimes happens that there are two. This is the case of /y/, to be seen in (2), which yields either /i/ or /u/ in several languages, with one strategy or the other usually being clearly dominant within a language. Yet both adaptations are phonologically grounded and minimal (/u/ results from delinking [–back], whereas /i/ from delinking [+round]). One might think that adaptations to /u/ result from orthography influence, since the grapheme <u> represents the vowel /u/ in many languages. However, a cross-linguistic perspective indicates that orthography provides a poor account and could not be responsible for so many adaptations to /u/. For example, orthography could not adequately account for the 49.2% of adaptations to /u/ in Moroccan Arabic and the 97.2% in Russian, since the grapheme <u> does not exist in either the Arabic or the Cyrillic writing systems, and is thus meaningless in those systems. In fact, <u> generally corresponds to a sound other than /u/ in L2. Even for English, which adapts French /y/ (<u>) to /u/ in 89.7% of the cases, orthography provides an unlikely explanation. L2 (French) orthography influence cannot be held responsible, because the French vowel /u/ is represented by the digraph <ou>, not the simple grapheme <u> (e.g. French *sou* [su] 'penny') and, in English, <u> represents /ʌ/ 63% of the time, /u/ 10% of the time, and /u/ only 2% of the time (Birch 2002: 85, based on Venezky's 1970 study of the sound-spelling correspondences in a corpus of the 20,000 most common English words, following the pronunciations provided by Kenyon and Knott 1953). Thus, knowledge of English orthography would not greatly incline English borrowers to interpret <u> (/y/) in French words as /u/. To sum up, adaptation seems to be overwhelmingly phonological and the influence of orthography very low in the database, either among malformations or outside malformations, as we will show in (3).

4.2 Targeted Corpora and an Overview of the Segmental Changes Directory

The CoPho loanword database also includes several targeted corpora, which were assembled to clarify a given point, address a specific question, or to test a particular hypothesis. In targeted corpora, as opposed to general ones, all the forms are pertinent to the particular enquiry, for example all containing a particular sound. These corpora are of varying size, depending on their purpose. Among them is one on proper names (more specifically brand names), whose general purpose is to explore the potentially marginal status of proper names in loanword adaptation, and to more thoroughly verify the influence of orthography.

An investigation of the atypical deletion of /h/ in English loanwords in French led to a consideration of the treatment of gutturals more generally and to the building of two corpora of guttural-containing Arabic loanwords introduced into French and English, as well as several small corpora of English borrowings, which all include the guttural /h/, in Greek, Bulgarian, Catalan, Russian, Mandarin Chinese, and so on. These corpora are described in Paradis, LaCharité, and Brault (1999) and Paradis and LaCharité (2001). Project CoPho also has a large database of French, Turkish, and German borrowings containing the front rounded vowels /y/ and /ø/ in Russian, and a similar corpus of French borrowings in Khmer. The objective of these corpora was to check whether vowels other than nasal vowels systematically unpack in L1s that do not allow them. This question is addressed in Paradis and Prunet (2000) and further developed in Paradis and Thibeault (2004) and Paradis (2006). More recently, in order to verify whether the phonetic aspiration of stops in English has an impact on their categorization in Mandarin Chinese, where stop aspiration is distinctive, a large targeted corpus of English borrowings in Mandarin Chinese was constructed. The details and analysis of this corpus are reported in Paradis and Tremblay (2009). Several corpora intended to verify specific predictions of the perceptual stance were also created under the auspices of the project. Some of these are described and discussed at length in LaCharité and Paradis (2005).

For all the corpora, segmental adaptations have been compiled into a directory listing the adaptation(s) for each ill-formed segment. An example of the tables found in the directory for the vowel /y/ is shown in (2).

(2) Adaptations of French /y/ in Project CoPho's general and targeted corpora

Language	*Adaptation to /i/ ([I])*	*Adaptation to /u/ ([U])*
Canadian English	–	70/78 (89.7%)[5]
Fula	56/56 (100%)	–
Kinyarwanda	121/167 (72.5%)	46/167 (27.5%)
Lama	3/31 (9.7%)	28/31 (90.3%)
Lingala	142/155 (91.6%)	13/155 (8.4%)
Montreal Italian	–	22/22 (100%)
Moroccan Arabic	68/134 (50.8%)	66/134 (49.2%)
Russian	10/358 (2.8%)	348/358 (97.2%)[6]
Spanish	1/20 (5%)	19/20 (95%)

The directory allows rapid visualization of the most frequent adaptation(s). Its immediate purpose is to enable us to accurately quantify the adaptations and other modifications that occur in loanword adaptation. Its ultimate objective is to allow us to make more precise predictions regarding segment adaptation, in the light of the phonologically principled adaptation strategies available in each language, and to also verify whether some pieces of information (specific features, class nodes, etc.) are more resistant to change than others, and if so, which ones.

5 Major Findings

5.1 General Statistics and Findings

The main findings yielded by the analysis of the general corpora of Project CoPho are summarized in the table in (3). As shown and as will be addressed below, L1 phonological constraints can account for the vast majority of the sound modifications that we observe in the 29,355 borrowing forms of the Project CoPho database. A borrowing form is the pronunciation of a borrowing by a native speaker consultant. The number of borrowing forms is thus usually greater than the number of borrowings (12,452). Malformations are computed on the basis of borrowing forms, which, unlike borrowings, are primary data. Repair of malformations is overwhelmingly phonological (50,091/54,443 malformations, i.e. 92%). Non-phonological cases (those that cannot be straightforwardly accounted for by the insertion or deletion repairs that apply to satisfy the L1 phonological constraints) represent only 4,351/54,443 of the malformations (8%). As will be shown in Section 5.2, non-phonological sound modifications are due mainly to analogy or orthography influence.

(3) Project CoPho's general statistics including segmental, syllabic, and stress malformations from the 13 corpora of French and English loanwords (updated July 29, 2009)

General information			*Malformations*					
Corpora	*Loans*	*Forms*	*Total*	*Phonological cases*[7]				*Non-phonological cases*
				Total	*Adaptations*[8]	*Importations*	*Deletions*	*Total*
English borrowings in								
Old Quebec French	485	597	489	397 81.2%	298 75.1%	77 19.4%	22 5.5%	91 18.6%
Parisian French	901	2,576	3,153	2,749 87.2%	1,570 57.1%	987 35.9%	192 7%	404 12.8%
Quebec City French	949	2,416	2,434	2,183 89.7%	1,479 67.7%	602 27.6%	102 4.7%	251 10.3%
Montreal French	949	2,248	2,285	2,099 91.9%	1,262 60.1%	747 35.6%	90 4.3%	186 8.1%
Mexican Spanish I	1,045	1,514	3,137	3,008 95.9%	1,583 52.6%	1,317 43.8%	108 3.6%	129 4.1%
Mexican Spanish II	1,034	2,342	5,645	4,490 79.5%	2,836 63.2%	1,569 34.9%	85 1.9%	1,155 20.5%
Japanese	1,167	2,991	7,760	7,373 95%	6,778 91.9%	492 6.7%	103 1.4%	387 5%
Calabrese Italian	2,161	5,191	14,740	14,438 98%	6,182 42.8%	7,821 54.2%	435 3%	302 2%

French borrowings in								
Canadian English	674	1,667	1,034	748 72.3%	555 74.2%	137 18.3%	56 7.5%	286 27.7%
Moroccan Arabic	1,127	2,685	4,275	3,979 93.1%	3,104 78%	568 14.3%	307 7.7%	296 6.9%
Kinyarwanda	756	2,130	4,639	4,207 90.7%	4,119 97.9%	26 0.6%	62 1.5%	432 9.3%
Lingala	672	1,917	3,734	3,408 91.3%	3,396 99.6%	2 0.1%	10 0.3%	326 8.7%
Fula	532	1,081	1,118	1,012 90.5%	908 89.7%	45 4.5%	59 5.8%	106 9.5%
total for all corpora	12,452	29,355	54,443	50,091 92%	34,070 68%	14,390 28.7%	1,631 3.3%	4,351 8%

Phonological adaptation, which is called simply "adaptation" here, is the modification or replacement (i.e. repair) of an L2 sound or structure to comply with one or more L1 phonological constraints. Adaptation is geared to ensuring that the L1 system remains unchanged (see LaCharité and Paradis 2005 for discussion), and it is the norm. The figures in (3) also show that adaptation of a foreign phonological entity, as opposed to its deletion, is the norm. In fact, deletion is rare, occurring in only 1,631/50,091 (3.3%) phonological cases. It is well below 10% in all the corpora. We attribute this to the Preservation Principle in (4) (see, e.g. Paradis, Lebel, and LaCharité 1994; Paradis and LaCharité 1997; and Paradis and LaCharité 2011).

(4) Preservation Principle:

Segmental information is maximally preserved, within the limits of constraint conflicts.

When foreign sounds are not adapted or (much more rarely) deleted, they are imported, that is, left unadapted. Non-adaptations account for 14,390/50,091 (28.7%) of the phonological cases. The more bilinguals there are in a community, the more non-adaptations we find. Thus, there are more non-adaptations in the corpus of Montreal French (35.6%) than there are in the corpus of Quebec City French (27.6%). According to Statistics Canada's figures for the period circa 1996 (www.statcan.ca), which corresponds to the period that QF both corpora were constructed, the rate of bilingualism in Montreal was 49.7%, and that of Quebec City was 30%. The correlation between a high rate of bilingualism and a high rate of importation is not surprising since, as pointed out by Danesi (1985: 18–19), borrowers often attempt to imitate the pronunciation of the L2 words to the best of their abilities. We call this "intentional phonetic approximation." We hypothesize that the poor imitations of individual speakers are retouched by the most bilingual speakers of a community in the same way that poor adaptations are corrected

before leaving the bilingual circle (see Grosjean 1982 on this question). Crucially, intentional phonetic approximation (importation) is an attempt to have the L1 phonological system accommodate characteristics of L2. Therefore, it can introduce L2 sounds and structures into L1. In contrast, naïve phonetic approximation, which is the kind to which supporters of the phonetic approximation stance generally refer, cannot. Naïve phonetic approximation occurs precisely because the borrower (a monolingual) fails to perceive foreign sounds or structures accurately. Naïve phonetic approximation perceptually warps malformations towards sounds and structures familiar to the L1.

5.2 Findings Concerning the Influence of Orthography

Adaptations are usually systematic, minimal, and sound based. However, there are occasionally changes in borrowings that are not sound based, but are instead based on orthography. As already mentioned, this influence is weak in the CoPho database as a whole (959/54,443 malformations, 1.8%). Even if we exclude stress and metrical malformations and consider only segmental and syllabic ones, as in the table in (5), the percentage remains very low (959/48,621 malformations, 2%).[9] This is consistent with findings in lexical processing that "script does not provide a basis for lexical representation" (Brown *et al.* 1984: 501).

(5) Statistics on orthography influence in the segmental and syllabic malformations of the Project CoPho loanword database

Corpora	*Segmental*	*Syllabic*	*Total*
of English borrowings	441/22,958 1.9%	81/10,863 0.7%	522/33,821 1.5%
of French borrowings	171/9,339 1.8%	266/5,461 4.9%	437/14,800 3%
total	612/32,297 1.9%	347/16,324 2.1%	959/48,621 2%

For reasons that we have not yet explored, the influence of orthography seems, on average, slightly higher in the French borrowings of the CoPho database than in the English ones, though spelling-driven adaptations are far from being usual in either case. Danesi (1985: 22) reports similar results, saying that "orthographically-induced nativization patterns" are non-existent in his corpus of approximately 500 English loanwords in Canadian Italian. This is not surprising since, as consistently indicated by sociolinguistic studies (e.g. Field 2002),

loanwords, apart perhaps from some proper names, are borrowed in their oral form the great majority of the time.

As the following table shows, the average of orthography influence in the adaptation of ill-formed segments in all 13 of Project CoPho's corpora is under 5%. In fact, orthography influence exceeds this level only in the Parisian French corpus (5.3%) and is considerably lower than 5% in several other corpora. In the Moroccan Arabic corpus, orthography influence is non-existent, whereas in Montreal French and Japanese, it is only 0.7% and 1%, respectively. It is further worth noting that orthography influence, on whole, is almost as weak among vowel malformations, 2.1% (518/24,414), as among consonant malformations, 1.5% (92/6,285), as further indicated in (6).

(6) Statistics on the per language orthography influence among vowel and consonant malformations of the Project CoPho loanword database

Corpora	*Vowels*	*Consonants*	*Total*
English borrowings in			
Old Quebec French	2/305 0.7%	0/96 0%	2/401 0.5%
Parisian French	115/1,892 6.1%	28/805 3.5%	143/2,697 5.3%
Quebec City French	22/1,464 1.5%	2/613 0.3%	24/2,077 1.1%
Montreal French	13/1,363 0.9%	0/576 0%	13/1,939 0.7%
Mexican Spanish I	20/1,221 1.6%	4/597 0.7%	24/1,818 1.3%
Mexican Spanish II	75/2,187 3.4%	15/868 1.7%	90/3,055 2.9%
Japanese	38/2,962 1.3%	0/773 0%	38/3,735 1%
Calabrese Italian	73/5,840 1.3%	34/1,396 2.4%	107/7,236 1.5%

Corpora	*Vowels*	*Consonants*	*Total*
French borrowings in			
Canadian English	7/678 1%	0/65 0%	7/743 0.9%
Moroccan Arabic	0/3,033 0%	0/736 0%	0/3,769 0%
Kinyarwanda	81/2,026 4%	0/14 0%	81/2,040 4%
Lingala	41/908 4.5%	3/1,033 0.3%	44/1,941 2.3%
Fula	31/535 5.8%	8/311 2.6%	39/846 4.6%
Total	518/24,414 2.1%	92/6,285 1.5%	610/30,699 2%

Ill-formed sounds are not the only ones that could be influenced by orthography; if orthography were really a dominant factor in loanword adaptation, one might reasonably expect it to influence the interpretation of well-formed sounds, that is, those that occur in both L1 and L2, as well as malformations. However, we found that spelling-driven modifications are as rare among well-formed sounds as among the malformations in the database. To determine this, we examined discrepancies between an L2 grapheme and the sound it normally represents in the Latin alphabet or in the L1 writing system. For example, the grapheme <u> represents /u/ in the majority of languages that use the Latin alphabet, including Italian, Spanish, Lingala, Fula, and Kinyarwanda (though in French <u> stands for /y/). As another example, unlike most other languages using the Latin alphabet, where <i> corresponds to /i/, /i/ is much more often represented with <e> in English (*even* [ivən]) than with <i>. We also considered that orthography could lead to the realization of silent L2 phonemes and letters (e.g. French *franc* [fʀɑ̃] that yields MA [fran̲k], not [fran], based on the grapheme <c> in *franc* which is often pronounced (as /k/) in French, as in *sac* [sak] 'bag', *bloc* [blɔk] 'block', etc.). Though we have not yet studied the influence of orthography outside malformations for all corpora, for those that we have examined, it is clearly negligible. For instance, in the Moroccan Arabic (MA) corpus, orthography influence affects only 0.5% (14/2,685) of the forms for which there is at least one discrepancy of the sort exemplified above. In Fula, the rate of orthography influence outside malformations is almost as low; only 2% (22/1,081) forms display a non-transparent grapho-phonemic correspondence, that is, a phoneme that is represented by an L2 grapheme or a multigraph that is different from the IPA or the IPA-like symbol and that could potentially be misleading with respect to its pronunciation.

Lamoureux (2000) showed that the influence of orthography in the Canadian English (CE) corpus was also very low. The author considered many spellings that might yield grapho-phonemic confusion due mainly to differences between the spelling conventions of French and those of English, or to divergences between the L2's conventions and those that are cross-linguistically more widespread. Among the spellings considered were French mute <e>, <h> and <t>, the simple grapheme <s> when pronounced [z], the digraphs <ai> [ɛ], <au> [o], <ch> [ʃ], <eu> [ə], <gu> [g], <ll> [j], <oi> [wa], <ou> [u], <ph> [f], <qu> [k], <th> [t] and the trigraphs <eau> [o], <eoi> [wa] and <oeu> [ə] (e.g. Fr. *blouse* [bluz] 'blouse' → CE [blaʊz] because the digraph <ou>, which represents /u/ in French, most often refers to the diphthong /aʊ/ in English). Yet despite the fact that the French grapho-phonemic correspondences are frequently not those to which English speakers are accustomed, confusions are rare. Lamoureux reports that orthography influence can account for only 93/2,038 (4.6%) of the cases studied.

Goulet (2001) studied the English simple graphemes <c>, <s>, <g>, <a>, <u>, <e>, <o>, <i> in the Japanese corpus, because these graphemes represent a variety of pronunciations in English. For instance, <c> can represent /k, s, ʃ, tʃ/, <s>, /s, z, ʃ, ʒ/, <g>, /g, dʒ, ʒ/, <a>, /æ, ə, e, ɑ, ɛ/, <u>, /ʌ, ʊ, ju, w, ə, u, ɪ/, <e>, /ɛ, ə, i, ɪ/, <o>, /ɑ, u, o, ə, ʌ, ʊ, ɪ/, and <i> /ɪ, aɪ, ə, i, ʌ, j/. In some cases these graphemes can even be mute (e.g. <c> and <g> in English *scene* [sin] and *sign* [saɪn], respectively). In spite of these variations, which could be extremely misleading for a non-native speaker, the rate of orthography influence outside malformations in this corpus is only 2.4% (148/6,095 cases).

Fisher (2004) reported similar results for the corpus of English loans in the Calabrese Italian (CI) corpus. The influence of orthography outside malformations was evident in only 191/20,144 cases (0.9%). Fisher's methodology is different from that used by Lamoureux (2000) and Goulet (2001); instead of establishing his classification on the basis of graphemes, it is done on the basis of phonemes. The English phonemes considered are /e/, /i/, /o/, /aɪ/, /aʊ/, /d/, /dʒ/, /k/, /l/, /m, n/, /p/, /s/, /t/, /tʃ/, /ʃ/, /v/, /z/. We do not favor this methodology because not all these phonemes are represented in an equally non-transparent way in English. For instance, /s/ is usually represented by <s> (which is pronounced as such by CI consultants); <s> does not usually stand for /z/, nor does <z> usually represent /s/. Although this sometimes occurs, and although single consonants are sometimes written with a double letter in English (e.g. <pp>, <ll>, <kk>, <ss>), a representation that sometimes misleadingly yields a phonemic geminate in CI (e.g. Eng. *pussy cat* [pʊsikæt] → CI [pussikat]), the sound-spelling correspondences for consonants are generally more transparent than those for vowels. The English phonemes in Fisher's study that constitute clearer values for the purpose of measuring the influence of orthography are /e/, /i/, /aɪ/ and /aʊ/, which are represented variously in English and whose representations never conform to the usual conventions of the Latin alphabet. For instance, /e/ is represented with <a>, <ai>, <ea>, not as more usual and transparent (closer to the IPA system) <e>; /i/ is spelled with <e>, <ee>, <ea>, <ey>, much more often than with <i>; /aɪ/ is spelled as <i>, <y>, <ie>, rarely as <aj>; /aʊ/ is usually spelled

as <ou>, <ow>, not as <aw>. Yet the rate of orthography influence (e.g. Eng. *scr͡aper* [skɹepəɹ] → CI [skra͡pər]) is almost as low as when all the phonemes studied by Fisher are considered: 27/2,567 cases (1%).

Beaulieu and Project CoPho's (2004) study of English graphemes in the Quebec City French (QCF) corpus of loanwords, which includes 44 simple graphemes, digraphs; and trigraphs, also reveals a very low rate of orthography influence: 86/9,909 cases (0.9%). Even if we restrict ourselves to the vowels selected from Fisher's study (except for /a͡ʊ/, which is not permitted in French), the rate is still as low: 7/759 (0.9%); for example, Eng. *tux͡edo* [təksido] → QCF [tɔkse͡do]). This again indicates that adapters very rarely rely on orthography to establish their pronunciation of loanwords, in spite of marked grapho-phonemic discrepancies between the L1 and L2 writing systems, or between the L1 and L2 writing conventions and those more generally used by the other languages written in the Latin alphabet.

Much to our surprise, this also proved true of proper nouns, whose pronunciation we expected to be supported by orthography to a greater extent than common nouns because it seemed to us that proper nouns, especially brand names, are seen in their written form much more often. Also, even if bilinguals are the online borrowers and adapters generally speaking, monolinguals, too, might need to adapt proper nouns. Yet, according to Vincent and Project CoPho's (2004) study of 43 English brand names, such as *Ashton*, *Burger King*, *Canadian Tire*, *Price Club*, and so on, which were pronounced by six French-speaking consultants from, and living in, Quebec City, this is not the case. The 29 graphemes examined in the 163 forms of Vincent and Project CoPho's corpus undergo the influence of orthography in only 5/266 cases (1.9%). For instance, *Febreeze* [fɛbɹiz] → QCF [fibɹiz], where <e> is taken to follow the more general English pattern and is erroneously pronounced /i/ in QCF, and *Lysol* [la͡ɪsɑl] → QCF [lisɔl] because <y> stands for /i/ in French, never for /a͡ɪ/ as in English). Among proper nouns, analogy plays a greater role, affecting 29/266 cases (10.9%). For example, *Canadian Tire* [kənediənta͡ɪəɹ] → QCF [kanadjɛnta͡ɪəɹ] based on French *Canadienne* [kanadjɛn] 'Canadian (fem.)', and *Winners* [wɪnəɹz] → QCF [wenəʀ] where English /əɹ/ is interpreted as the French agentive *-eur* /-əʀ/. Another surprise was the extraordinarily high rate of importations in the targeted corpus of proper nouns – it reached 42.5% (91/214). This contrasts sharply with the 27.6% rate of importations observed in (3) for the general corpus of QCF. Perhaps native speakers do not feel as free to adapt proper nouns as they do common nouns (this was probably truer in 2004 than in 1996, since bilingualism is now greater in Quebec City than it was in 1996, when the Quebec French corpora were assembled). When it comes to the pronunciation of people's names or brand names, speakers may want their pronunciation to be as faithful as possible to the source language form, perhaps because proper nouns, as opposed to common ones, refer to entities that are unique (Molino 1982: 14–15).

Below is a summary table that recapitulates the figures provided above. Even if we cannot add the numbers and percentages because some refer to cases and others to forms (studies based on cases are more precise and, now favored by

Project CoPho), it is clear that the influence of orthography outside malformations is very low.

(7) Summary of the influence of orthography outside malformations in the Project CoPho corpora for which this information is available

Corpora	*%*	*Number of cases*
French loans in Moroccan Arabic	0.5%	14/2,685 forms
French loans in Fula	2%	22/1,081 forms
French loans in Canadian English	4.6%	93/2,038 cases studied
English loans in Calabrese Italian	0.9%	191/20,144 cases studied
English loans in Quebec City French	0.9%	86/9,909 cases studied
English brand names in Quebec City French	1.9%	5/266 cases studied

The corpus for which the rate of orthography influence appears to be the greatest is CE but many cases in this corpus are not clear-cut; they could also be due to analogy influence. For instance, <ill> in French represents /j/ but it is systematically interpreted (22/22 cases) as /l/ in English (e.g. Fr. *Chantilly* [ʃãtiji] and *bouillon* [bujõ] → CE [ʃæntɪli] and [buljən]). Although orthography provides one possible explanation, we cannot rule out analogical influence based on older French borrowings, where <ill> was indeed pronounced as [il] in French (*bouillon* stems from Fr. *bulle* [byl] 'bubble', which meant literally 'to cook with bubbles'). We must not forget that French borrowings began to be introduced into English a thousand years ago and the amount of borrowing from French was massive in the fourteenth and fifteenth centuries. Until the seventeenth and even the eighteenth century, French still had the palatal lateral /ʎ/, which was graphically represented by <ill> and which subsequently became /j/ in the language (Harris 1992). The phonological adaptation from the French /ʎ/ would have been to English /l/. Even if the French borrowings of the CE corpus are no older than 200 years, some of the borrowings include older borrowings. For example, this is the case of late nineteenth century *brillantine* [bɹɪljəntɪn], from French [bʀijãtin], which is very likely adapted by analogy to *brilliant* [bɹɪljənt] (borrowed from French *brillant* in the seventeenth century) and of mid-twentieth-century *dérailleur* [diɹeləɹ], from French [deʀajəʀ], adapted most likely on the basis of English *rail* (borrowed from French *reille* in the sixteenth century). In other words, the pronunciation [l] for [j] in the above examples might result from analogy to older French pronunciation, when /Vil/ was still pronounced [Vil] instead of current [Vj].

However, even though orthography's influence is slight in the database, overall, it does account for a notable proportion of the non-phonological modifications

of segmental and syllabic malformations, 959/3,827 cases (25.1%). Among French borrowings, it accounts for 437/1,446 (30.2%) of the non-phonological cases; among the English ones it is a bit weaker, accounting for 522/2,381 (21.9%) of the non-phonological cases.

(8) Orthography influence in the segmental and syllabic malformations of the Project CoPho loanword database/number of non-phonogical modifications

Corpora	*Segmental*	*Syllabic*	*Total*
of English borrowings	441/1,826 24.2%	81/555 14.6%	522/2,381 21.9%
of French borrowings	171/792 21.6%	266/654 40.7%	437/1,446 30.2%
Total	612/2,618 23.4%	347/1,209 28.7%	959/3,827 25.1%

Analogy is the second most influential non-phonological factor. Although we have not yet compiled tables of statistics for all the non-phonological factors in all 13 corpora, this has been done for the Parisian French (PF) and the Old Quebec French (OQF) corpora. The statistics for the Parisian corpus are presented below.

(9) Statistics on non-phonological factors in Parisian French for segments and syllables

Non-phonological cases	*Ortho-graphy*	*Analogy*	*Variable pronunci-ation*	*Phonetic approx-imation*	*Partial analogy*	*Missed targets*	*False analogy*	*Remote forms*	*Hyper-correction*
404	169	7.3	62	48	25	16	5	4	1
12.8%	41.8%	18.1%	15.3%	11.9%	6.2%	4%	1.2%	1%	0.2%

"Non-phonological" is used here as a cover term to include many different types of cases, some of which are not necessarily inimical to the phonological position. Rather, some such cases simply provide no insight into the sound-related changes in loanword adaptation and must therefore be disregarded. This is the case of variable pronunciations that account for 62/404 (15.3%) of the non-phonological cases in PF. This category refers to borrowed words whose phonemic content in English is variable and whose input form cannot be reliably ascertained. For example, *ketchup* may be pronounced [kɛtʃəp] or [kætʃəp], so we cannot be sure whether the input form contained /ɛ/ or /æ/. Nor are missed targets, which constitute 16/404 (4%) of the non-phonological cases in PF, revealing. For the most part, these are L2 words that are truncated in L1 and whose truncated part contains the malformation that we wish to track (e.g. PF *zap* for English *zapping*, where the suffix *-ing* that contains the foreign phoneme /ŋ/ was omitted). The

4/404 (1%) remote forms are those for which we cannot provide an explanation in terms of phonology, phonetics, analogy, orthography, and so on (e.g. PF/ɑ̃ŋgəs/ for English *haggis*), whereas hypercorrection (1/404, 0.2%) is an L1 phenomenon. The influence of these last three factors is always very small. Compare the influence of orthography (41.8%) or even true analogy (18.1%) to other types of non-phonological factors.

Nonetheless, among non-phonological factors, the influence of orthography is generally not as high as in the PF corpus. For instance, in OQF, even if the rate of non-phonological cases is among the highest in the Project CoPho loanword database, at 18.6% (91/489), the influence of orthography is low, 7.7% of the non-phonological cases (7/91), as shown in (10).

(10) Statistics on non-phonological factors in Old Quebec French for segments and syllables

Non-phonological cases	*False analogy*	*Phonetic approx-imation*	*Variable pronunci-ation*	*Remote forms*	*Partial analogy*	*Ortho-graphy*	*Missed targets*	*Analogy*	*Hyper-correction*
91 (/489)	25	24	16	7	7	7	5	0	0
18.6%	27.5%	26.3%	17.6%	7.7%	7.7%	7.7%	5.5%	0%	0%

The most influential non-phonological factors in the OQF corpus are phonetic approximation (24/91, 26.3%) and false analogy (25/91, 27.5%). The atypically high rate of non-phonological influence is not surprising since the majority of the population in Quebec at the time the borrowings of this corpus were collected (i.e. the late nineteenth and early twentieth centuries), was monolingual. Only 8% of the population lived in bilingual areas such as cities. Meaning is central to adult lexical processing (Wingfield 1978: 180; Gleason and Ratner 1998: 168; Holliday and Weekes 2006), so trying to make sense of the words they encounter of course affects the interpretation of borrowings by monolinguals, who have recourse to the semantic associations of L1 only. A high rate of false analogy is thus understandable and explains how, for example, English *murder* ended up being interpreted as *mordeur* 'one who bites' in OQF, based on Fr. *mordre* 'to bite' (for more examples and a thorough discussion, see Paradis and LaCharité 2008).

5.3 Findings Concerning Phonological Patterns

Our phonological bias in the study of loanword adaptation is supported by several arguments that are discussed thoroughly in LaCharité and Paradis (2005). Here we will focus on two of these arguments, the stability of adaptation and monotonicity in adaptation, and address some of their apparent counterexamples.

5.3.1 Stability of Adaptation and Discussion of Apparent Counter-examples

Distinguishing clearly between usual and exceptional adaptations reveals a remarkable stability of adaptation patterns both within and across languages. For

example, in the adaptation of English loanwords, /ɪ/ is overwhelmingly adapted as /i/ in the corpora of Calabrese Italian, Canadian and Parisian French, Mexican Spanish, and Japanese. This is true whether we consider each individual language or all four languages collectively. Strong stability patterns suggested to us that loanword adaptation must be mainly phonological, not phonetic, because, phonetically, each language's /i/ is distinct, meaning that the formant values of /i/ are different from language to language. English /ɪ/, which occurs in a variety of phonetic environments in loanwords, is not phonetically closest to the instantiations of /i/ in all the languages we have studied. Cross-linguistic differences in the formant values for a particular vowel phoneme, along with differences in vowel inventories, also mean that the acoustic territory occupied by a particular L2 vowel phoneme may overlap with that of a different L1 vowel. For example, English /ɪ/, intrudes on the acoustic space occupied by French /ɛ/ (see Delattre 1981; Martin 2002 on the differences between English and French). As in the case of English [b] that corresponds to French [p] in terms of VOT, but which is nonetheless treated as /b/ in French, this case, in which English /ɪ/ is adapted as French /i/, not /ɛ/, illustrates the general point that adaptation is based on a sound's phonological status (i.e. category), not its phonetic realization. Stability of adaptation patterns within and across languages is thus one of the findings that argue against the phonetic approximation analysis and support our phonological bias.

Nonetheless, there are cases of variation (apparent lack of stability) that might, at first glance, be interpreted as cases of phonetic approximation. Let us consider the classic case of interdentals, which has routinely been interpreted as a counterexample to stability, because of interlinguistic variation. Intralinguistically, interdental adaptation is quite stable (see Hyman 1975 on this issue). The English interdental fricatives /θ/ and /ð/ are adapted as /t/ and /d/, respectively, in most languages. However, as is well known, European French adapts the interdentals as /s/ and /z/ in loanwords (see Hyman 1975) as do Japanese and German. We initially considered that there were two phonologically possible adaptations (/θ/ → /t/ or /s/; /ð/ → /d/ or /z/), or that there was one phonological adaptation (to /t/ and /d/) and one due to perceptual confusion (to /s/ and /z/). However, we rejected that possibility, or that both adaptation patterns were perceptually driven, because it did not square with what is known about the acoustics of the sounds involved, nor the confusions often seen in first and second language acquisition. Indeed, we have generally found that adaptations in loanwords do not reflect the perceptual behavior and biases of monolingual L1 speakers or bilinguals with a low level of proficiency (see LaCharité and Paradis 2005), which constitutes another reason for our phonological bias. In the interdental case, if we consider main resonant frequencies, attested perceptual confusions, type of sound spectrum, and amplitude, the best perceptual match for [θ] is [f], not [s] and certainly not [t], as discussed in depth in Paradis and LaCharité (2002).

Instead, we came to view the adaptation of [θ] to [t] and of [ð] to [d] as the phonological adaptations of these sounds and the /s/, /z/ realizations as flawed

importations, meaning that they are not completely successful attempts to pronounce the English interdentals. These attempts are what we call in LaCharité and Paradis (2005) intentional phonetic approximation, as opposed to naïve phonetic approximation, except that in this case, intentional phonetic approximation is imperfectly performed. This would be in line with research by Davidson (2007) showing that English speakers attempting to produce Slavic clusters such as /vb/ that are not permitted in English produced a transitional (not a lexical) schwa between the consonants and that monolinguals were subsequently forced to fit these productions into the L1 system. In European French, which formerly did adapt the interdental fricatives to /t/ and /d/ (Fouché 1959: 369), the flawed importation to [s] and [z] has become institutionalized, meaning that it is taught in schools as a more prestigious and educated pronunciation of the English interdentals than [t] and [d].

Arabic provides another example of social constraints that have had an impact on the adaptation of the foreign sounds /θ/ and /ð/. Carter (2001: 24) reports that Classical Arabic /θ/ and /ð/ are adapted as stops in everyday words introduced into dialects of Arabic, but as /s/ and /z/ in more religious and literary words. We currently hypothesize that social factors underlie the realization of /θ/ and /ð/ as /s/ and /z/ cross-linguistically and that "the [s/z] realization is typical of communities where the source language . . . is perceived as prestigious, elegant and generally 'non-threatening'" (Paradis and LaCharité 2002).

The social context of loanword adaptation is thus another important factor that has to be considered and that can explain deviation from stability in loanword adaptation. One must resist the temptation to jump to a conclusion regarding loanword adaptation, without considering the broader context in which it occurs, that is, without knowing as much as possible about various facets of the borrowing and source languages and the social and historical contexts in which borrowing occurs, because these factors can have an impact even at the phonological level. Phonology is certainly central to the topic, but we also have to delve into L2 acquisition, sociolinguistics, and history, to name just a few of the areas that borrowing routinely intersects. There are several types of bilingualism (individual, institutional, collective, and so on; see Mackey 1992: 39) and borrowing is often an emotionally charged social issue depending on the type of bilingualism that is practiced in a community. A borrower that holds the L2 in high esteem will be more likely to import or try to import a foreign sound than a borrower who has strong feelings against the L2. As an aside, importations are also of particular interest because they indicate that borrowers must generally perceive foreign sounds and structures correctly, which is another reason for our phonological bias in classifying adaptations.

5.3.2 Monotonicity and Discussion of Apparent Counter-examples Monotonicity (one-to-one segment adaptation) in sound adaptation constitutes another finding that underlies our phonological bias. Monotonicity means that a single foreign sound systematically yields a single L1 segment, which is encoded by the Isomorphism Hypothesis, proposed by Paradis and Prunet (2000: 332). With a

couple of notable exceptions, we do not find a single foreign sound in the source language form corresponding to two sounds in the adapted form. The first exception is French nasal vowels, which unpack in languages that do not allow them. For example, the French words *coupon* [kupɔ̃)] and *ensemble* [ɑ̃sɑ̃bl] are adapted as [kupɑn] and [ɑnsɑmbəl] in English. Paradis and Prunet (2000) attribute unpacking to the phonological structure of the nasal vowel, which is comprised of two root nodes, one of which is unanchored to a consonant slot in French. English, like other languages that do not permit phonemic nasal vowels, repairs this ill-formed structure by delinking the nasal feature from the vowel and assigning its root node a consonant slot, yielding a vowel-nasal sequence. Unpacking in this case is a phonologically principled repair.

Other apparent cases of unpacking yield less easily to a phonological explanation and have been cited as evidence of phonetic approximation. This is the case of the front round vowel /y/ that often corresponds to /ju/ in borrowed words from French in English (recall French *butte* [byt] → English [bjut] in Section 2). The idea behind a phonetic approximation account would be that the properties of the illicit front round vowel are perceived, but as two distinct sounds, a front one and round one, since the combination in a single sound is prohibited in the L1. We have dispelled this explanation in Section 2.

The case of Russian, discussed at length in Paradis and Thibeault (2004) and Paradis (2006), might constitute another apparent counterexample to the Isomorphism Hypothesis. However, a closer examination disconfirms the idea that /y/ is perceptually unpacked in loanword adaptation and more generally confirms the principle of monotonicity. The crux of Paradis' (2006) argument is that /Cʲu/ sequences stemming from /Cy/ in Russian loanwords from French, German, and Turkish results from delinking [–back] from the disallowed front rounded vowel /y/ and subsequently relinking it to the preceding consonant, provided the Russian phonology permits the consonant in question to be palatalized (e.g. Fr. *étude* [etyd] 'study' and *lustre* [lystʀ] 'center light' → Russian [etʲúd] and [lʲústra], respectively; Turkish [tyk] 'bundle' → Russian [tʲúk]). Because the vowel /u/ is a non-palatalizing environment in Russian (palatalization is strictly restricted to the environment of front vowels in Russian), the facts are incompatible with a perceptual account and show that there is no perceptual unpacking of the foreign vowel /y/. In other words, /Cʲu/ sequences cannot result from phonetic approximation in Russian because they are perceptually very unnatural. Instead, palatalization results from the fact that loanword adaptation seeks to maintain as much phonological information as possible, meaning that the front round vowel's [–back] feature can be redeployed in just those cases where the preceding consonant is an appropriate phonological host. In situations where an appropriate host is unavailable, such as word initially, the feature [–back] is simply lost (e.g. French *unitaire* [ynitɛr] 'unitarian' → Russian [unʲitárnʲi] *[junʲitárnʲi]), even when the resulting form should be acceptable in the language.[10]

Khmer presents a reminiscent case. Khmer syllables have to be bimoraic, which means that a syllable must contain a long vowel, a diphthong, or it must be closed. To respect this requirement, in loanword adaptation, a short French vowel is

normally lengthened, for example, French *menu* [mə̲ny] is adapted as Khmer [mə̲ːnʊ̲j]. In contrast, in the case of /y/, diphthongization, not vowel lengthening, applies because, as in Russian, the front rounded vowel's [–back] feature is recuperated. Except that, unlike in Russian, [–back] is redeployed to the right, not to the left, to satisfy the language's bimoraic syllable requirement (e.g. French *ruban* 'ribbon' [ʀy̲bã] → Khmer [rʊ̲jɓaŋ] not *[rjʊ̲ɓaŋ]); see Paradis and Thibeault (2004: 181–182) for more details.

As mentioned earlier, phonological patterns do not emerge until a substantial number of pertinent loanwords are studied. In cases such as the Russian or the Khmer ones, this often necessitates the building of a targeted corpus and extensive research into the phonology of the borrowing language. Thus far, each time that we have undertaken an investigation of phenomena that are seemingly inimical to the phonological view, we have found a cogent phonological explanation for the facts. Of course, such research takes considerable time, so there are outstanding issues that remain to be addressed. One example is the apparent deletion of /r/ in codas, such as occurs consistently in the Japanese corpus (e.g. English *porch* [poɹtʃ] → Japanese [poːtʃi]), which is at first sight problematic for the Preservation Principle seen in (4). This case has been mentioned as a case of phonetic approximation, sometimes attributed to r-dropping that occurs phonetically in British English (e.g. Silverman 1992: 297). However, the general rarity of segment deletion in loanword adaptation leaves us skeptical of such an explanation. Perhaps, as in the case of h-deletion in English loanwords in French, Italian, Portuguese, and so on (discussed in depth in Paradis and LaCharité 2001), the answer lies in the phonological structure of /r/. Rhotics are prone to deletion cross-linguistically, especially those with a variety of phonetic realizations and in coda position, and they exhibit several phonological behaviors that are not yet well understood. Coda /r/s are often deleted, merged with, transformed into, or replaced by a vowel. To cite just a few of many possible examples, in German, where /r/ is phonetically uvular, coda /r/ can lower to something akin to a low vowel, so that *Tür* 'door' is realized as [tyːɐ] (Wiese 1996). In Quebec French, coda /r/, which is standardly realized as a uvular, is often deleted word-finally in informal speech (e.g. *bonjour* /bɔ̃ʒuʀ/ 'good day' → [bɔ̃ʒu(ː)], or in a cluster (e.g. *trois* 'three' /tʀwa/ → [twa]). During the Middle Ages /r/-deletion prevailed for such a long time in French that /r/ almost disappeared as a coda phoneme (Zinc 1986). The deletion of /r/ also applies in many Spanish dialects, especially in Caribbean Spanish. It could be that the rhotic is actually part of a diphthong, as proposed by several phonologists (see, e.g. Nikiema and Bhatt 2004 for their analysis of post-vocalic /r/-deletion in Haitian Creole).

Another issue that is, at first sight at least, problematic for the Preservation Principle and that we intend to explore further is word-initial vowel deletion that we find in the Moroccan Arabic corpus (e.g. Fr. *agrément* 'agreement' [agʀemã] → MA [grimɑ]; Fr. *autobus* 'bus' [ɔtɔbys] → MA [ṭʊbis]; Fr. *aspirateur* 'vacuum' [aspiʀatəʀ] → MA [spiratʊr]). The likelihood of this type of deletion is correlated with the number of syllables in the borrowing; the greater the number of syllables in the input form, the greater the likelihood that an initial vowel, which represents

an ill-formed syllable in MA as in many Semitic languages, will be deleted. Paradis and Béland (2002) report that in a bisyllabic vowel-initial word, the initial vowel is dropped in 25/126 cases (24%); in a trisyllabic word, initial vowel deletion occurs in 86/162 cases (53%); in words of four syllables, the initial vowel deletes in 26/40 cases (75%) and in words of five syllables, it deletes in 7/9 cases (78%). Given the rarity of deletion elsewhere in this corpus or, in any other corpus of the CoPho loanword database, or in the corpora of loanwords gathered by other phonologists (see, e.g. Danesi 1985; Ulrich 1997), and the fact that metrical constraints on words are well attested cross-linguistically and referred to as the "word length effect," we strongly suspect that there is a metrical constraint at work here, whose effect is augmented by the MA constraint against onsetless syllables. However, this issue remains to be investigated thoroughly.

5.4 From Borrowings to Borrowers: Selecting Consultants

Before ending, we want to say something about the borrowers and the selection of consultants. Working with first-hand data means gathering it directly from native speakers of the borrowing language. This is fieldwork, and as with fieldwork carried out for other linguistic purposes, identifying suitable consultants is crucial, but not always easy. For our purposes, we found that speakers who are too fluent in L2, the source language for borrowings, are not desirable consultants as they have an atypically high tolerance for foreign sounds and structures (i.e. non-adaptations or importations), especially in the data elicitation situation. As shown in (3), importations or non-adaptations occur in all the Project CoPho corpora to varying degrees and they do provide valuable insights into loanword adaptation. Paradis and LaCharité (2008) show the rate of importation to be correlated with the rate of community bilingualism. Importations are also interesting because they indicate that borrowers must generally perceive foreign sounds and structures correctly, which is, as already mentioned, another reason for our phonological bias in classifying adaptations. Still, in order to get a more representative picture of what the L1 does and does not tolerate phonologically, we found it preferable to select consultants who were either monolingual, or who had a low level of L2 proficiency. As far as possible, we used language consultants whose parents did not know the L2 either.

In fact, a high tolerance for foreign sounds and structures can eventually prevail within an entire community, if it has been heavily dominated by the L2. For instance, Noyer (2007) reports that among Mexico's Huave speakers, who are now virtually all fluent in Spanish, adaptations in Spanish borrowings are diminishing and Spanish phonetic processes are increasingly common, even in Huave words. In this case, the line between L1 and L2 phonology is increasingly blurred. In the early years of Project CoPho's research, we embarked on a study of Spanish loanwords in Guaraní, a language spoken primarily in Paraguay and Uruguay, only to have to abandon it because speakers of the L1 had virtually ceased adapting the Spanish sounds and structures. Because their behavior is phonologically less revealing, and possibly misrepresentative of loanword adaptation

cross-linguistically, bilingual consultants who are too fluent, or those from fluently bilingual societies, should be avoided.

There are also other considerations in the selection of consultants. There has to be a reasonable match between the consultants and the social and pragmatic domains in which the borrowings occur. For instance, many French loanwords in English come from the realms of gastronomy, décor, fashion, aircraft technology, and so on. In order to know those words, one must have a certain life experience that is unusual in young adults (such as university students). For the assembly of the corpus of French words in English we therefore found it preferable to recruit more mature adults whose interests and life experience made it much more likely that they would know the target vocabulary. Recall that we elicit words via their meanings, fill-in-the-blanks, and so on, so a word must be part of the consultant's lexicon in order for him or her to produce it. If one ignores this limitation and simply presents borrowings in written form, some consultants may try to be overly accommodating and guess at the pronunciation of words that they do not in fact know.

However, not all consultants try to be accommodating. For some speakers more than others, and for some languages more than others, the concept of borrowing words from another language is a very delicate topic that arouses strong feelings and raises defenses. Poplack *et al.* (1988: 76) report that social class membership is predictive of borrowing rates, so if a consultant consciously or unconsciously knows this and resists the idea of being identified with that particular social group, he or she may resist the idea that he or she exhibits similar behavior. We have found that some speakers strongly reject the idea that they or their language use borrowed words, particularly if native terms exist for the same or similar concepts. Thus, even if speakers with negative attitudes to borrowing can be convinced to participate in this type of research, they may not provide representative information.

In sum, gathering first-hand loanword data is fieldwork and one quickly discovers that not all native speakers make equally informative consultants for the purpose of loanword adaptation. For the building of a corpus, phonologists might find it preferable to focus on selecting a few consultants, who are not too fluent in the source language, but whose vocabularies are nonetheless likely to include a good proportion of the L2 words that have to be verified, and whose attitudes are neither too accommodating nor too normative.

6 Conclusion

Loanword adaptation has proven to be a particularly fruitful domain of inquiry. Even though Project CoPho has been working in the field for nearly 20 years now, exciting avenues for research continue to be discovered – a lifetime spent on the study of loanword adaptation will certainly not exhaust them. These avenues go from segment, syllabic, or metrical structure to the nature of processes (repair strategies) and their targets. The study of loanword adaptation informs us about

the nature of the phonological principles guiding those processes, hence, the nature of the human language capacity. Loanword adaptation also provides important insights into bilingual processing and the organization of linguistic information in the bilingual mind. In this chapter we have tried to share some of what we have learned about these issues and the methodology by which we have explored them. As in any area of science, different methodologies may lead to different results. This is why it is so important to be explicit about the methodology used in the study of loanword adaptation and to work with statistically based corpora and data. Perhaps apparently contradictory results can be reconciled when issues of methodology are taken into account.

NOTES

1 CoPho stands for Constraints in Phonology. The project is under the direction of Carole Paradis at Laval University in Quebec City.
2 Dohlus (2005) says that the adaptation of English /kæ/ and /gæ/ to Japanese /kja/ and /gja/ is phonetic approximation. Her argument is that the palatal lingual contact induced by the English front vowel /æ/ is perceived as the palatal glide by Japanese speakers, whose own language phonologically distinguishes /ka/, /ga/ from /kja/ and /gja/. Dohlus does not provide statistics, but implies that adaptation of /kæ/, /gæ/ to /kja/, /gja/ is the norm, while phonological adaptation to /ka/, /ga/ is unattested or at least much less frequent.
3 This adaptation might also stem from false analogy to French *plaque* 'plate'. Nevertheless, in this case, too, English /b/ is interpreted as French /p/.
4 Firm or likely borrowing dates are established on the basis of a variety of published sources.
5 The eight remaining adaptations are to /ʌ/. As already mentioned, yod insertion, as in *butte* [bjut] from French [byt], is not part of the adaptation process *per se* and is as unpredictable in French words introduced into English as in native English words.
6 Almost half of the Russian cases involve palatalization. French /y/ after a palatalizable consonant yields a /C^{j}u/ sequence in Russian (cf. Paradis and Thibeault 2004 and Paradis 2006 for a thorough discussion).
7 Percentages of phonological and non-phonological cases are calculated on the total number of malformations.
8 Percentages of adaptations, non-adaptations, and deletions are calculated on the total number of phonological cases.
9 We have included syllabic malformation statistics in (5), even though we do not address such malformations here, so that numbers across tables correspond.
10 Word-initial #/ju/ is permitted and relatively common in Russian (e.g. *juk* 'south' and *junga* 'ship's boy').

References

Abbi, Anvita. (1991) *Reduplication in South Asian Languages: An Areal, Typological and Historical Study*. New Delhi: Allied Publishers Limited.

Abboud, Peter. (1979) The verb in Northern Najdi Arabic. *Bulletin of the School of Oriental and African Studies* 42, 467–499.

Abd-El-Jawad, Hassan R., and Issam M. Abu-Salim. (1987) Slips of the tongue in Arabic and their theoretical implications. *Language Sciences* 9, 145–171.

Abdel-Massih, Ernest T. (1975) *An Introduction to Egyptian Arabic*. Ann Arbor, MI: University of Michigan.

Abercrombie, D. (1967) *Elements of General Phonetics*. Chicago: Aldine and Atherton.

Abu-Mansour, Mahasen Hasan. (1996) Voice as a privative feature, in *Perspectives on Arabic Linguistics VIII: Papers from the 8th Annual Symposium on Arabic Linguistics*, ed. by Mushira Eid, Amsterdam: Benjamins.

Adger, David. (2007) Stress and phasal syntax. *Linguistic Analysis* 33, 238–266.

Agnew, Arlene and Evelyn G. Pike. (1957) Phonemes of Ocaina (Huitoto). *International Journal of American Linguistics* 23, 24–27.

Ahn, Mee-Jin. (2000) *Phonetic and Functional Bases of Syllable Weight for Stress Assignment*. University of Illinois dissertation.

Ajíbóyè, Oladiipo and Douglas Pulleyblank. (2008) Nasal harmony in Mòbà. Manuscript, University of Lagos and University of British Columbia.

Akinlabi, Akinbiyi. (1985) *Tonal Underspecification and Yorùbá Tone*. University of Ibadan dissertation.

Akinlabi, Akinbiyi. (1997) Patterns of tonal transfer I. Presented ACAL 28, Cornell University.

Akinlabi, Akinbiyi and Mark Liberman. (2000) The tonal phonology of Yoruba clitics, in *Clitics in Phonology, Morphology and Syntax*, ed. by B. Gerlach and J. Grijzenhout, 31–84, Amsterdam: Benjamins.

Aksu-Koc, Ahyan and Dan I. Slobin. (1985) Acquisition of Turkish, in *The Crosslinguistic Study of Language Acquisition*, ed. by D. I. Slobin, volume 1, 839–878, Hillsdale, NJ: Lawrence Erlbaum Associates.

Al-Ani, S. H. (1970) *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague: Mouton.

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu

© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

Al-Masri, M. and A. Jongman. (2004) Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results, in *Proceedings of the 2003 Texas Linguistics Society Conference*, ed. by A. Agwuele, W. Warren, and S-H. Park, 96–106, Somerville, MA: Cascadilla Press.

Al-Mozainy, Hamza. (1981) *Vowel Alternations in a Bedouin Hijazi Arabic Dialect: Abstractness and Stress*. University of Texas, Austin dissertation.

Alber, Birgit. (1997) Quantity sensitivity as the result of constraint interaction, in *Phonology in Progress–Progress in Phonology, HIL Phonology Papers III*, ed. by Geert Booij and Jeroen van de Weijer, 1–45, The Hague: Holland Academic Graphics.

Alber, Birgit. (2005) Clash, lapse and directionality. *Natural Language and Linguistic Theory* 23, 485–542.

Albright, Adam. (2002) *The Identification of Bases in Morphological Paradigms*. UCLA dissertation.

Albright, Adam. (2007a) Gradient phonological acceptability as a grammatical effect. Manuscript, MIT.

Albright, Adam. (2007b) Natural classes are not enough: Biased generalization in novel onset clusters. Presented at the 15th Manchester Phonology Meeting, Manchester UK, May 24–26, 2007.

Albright, Adam. (2009) Feature-based generalisation as a source of gradient acceptability. *Phonology* 26, 9–41.

Albright, Adam, A. E. Andrade, and Bruce Hayes. (2001) Segmental environments of Spanish diphthongization, in *UCLA Working Papers in Linguistics, Number 7: Papers in Phonology 5*, ed. by Adam Albright and T. Cho, 117–151, UCLA.

Albright, Adam and Bruce Hayes. (2002) Modeling English past tense intuitions with minimal generalization, in *Proceedings of the 2002 Workshop on Morphological Learning, Association of Computational Linguistics*, ed. by Michael Maxwell, 58–69, Philadelphia, Association for Computational Linguistics.

Albright, Adam and Bruce Hayes. (2003) Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90, 119–161.

Albright, Adam and Bruce Hayes. (2006) Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations, in *Gradience in Grammar: Generative Perspectives*, ed. by G. Fanselow, C. Féry, M. Schlesewsky, and R. Vogel, 185–204, Oxford University Press.

Albro, Daniel M. (1998) Evaluation, implementation, and extension of Primitive Optimality Theory. Master's thesis, UCLA. Retrieved April 14, 2011 from http://www.linguistics.ucla.edu/people/grads/albro/ma.pdf.

Albro, Daniel M. (2005) *Studies in Computational Optimality Theory, with Special Reference to the Phonological System of Malgasy*. UCLA dissertation.

Alderete, John. (1999) *Morphologically Governed Accent in Optimality Theory*. University of Massachusetts, Amherst dissertation.

Alderete, John. (2001) Dominance effects as transderivational anti-faithfulness. *Phonology* 18, 201–253.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John J. McCarthy, and Suzanne Urbanczyk. (1999) Reduplication with fixed segmentism. *Linguistic Inquiry* 30, 327–364.

Alderete, John, Adrian Brasoveanu, Nazarre Merchant, Alan Prince, and Bruce Tesar. (2005) Contrast analysis aids the learning of phonological underlying forms, in *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. by John Alderete, Chung-hye Han, and Alexei Kochetov, 34–42, Somerville, MA: Cascadeilla Press.

Allen, G. D. and S. Hawkins. (1980) Phonological rhythm: Definition and

development, in *Child Phonology*, ed. by G. Yeni-Komshian, J. S. Kavanagh, and C. A. Ferguson, volume 1, 227–256, New York: Academic Press.
Allen, G. D. and Sarah Hawkins. (1978) The development of phonological rhythm, in *Syllables and Segments*, ed. by A. Bell and P. Hollien, 173–185, Amsterdam: North Holland.
Allen, G. D. and Sarah Hawkins. (1979) Trochaic rhythm in children's speech, in *Current Issues in the Phonetic Sciences*, ed. by H. Hollien and P. Hollien, 927–933, Amsterdam: Benjamins.
Allen, John, M. Sharon Hunnicutt, and D. Klatt. (1987) *From Text to Speech: The MITalk System*. Cambridge: Cambridge University Press.
Alshawi, Hiyan. (1992) *The Core Language Engine*. Cambridge, MA: MIT Press.
Altschuler, Daniel. (2009) Osage fills the gap: The quantity insensitive iamb and the typology of feet. *International Journal of American Linguistics* 75, 365–398.
Anceaux, Johannes Cornells. (1965) *The Nimboran Language: Phonology and Morphology*. 's-Gravenhage: M. Nijhoff.
Andersen, Torben. (1988) Consonant alternation in the verbal morphology of Päri. *Afrika und Übersee* 71, 63–113.
Anderson, Anne, Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry Thompson and Regina Weinert. (1992) The IICRC Map Task corpus. *Language and Speech* 34, 351–366.
Anderson, Gregory and K. David Harrison. (1999) *Tyvan*. Languages of the World/Materials. Munich: LINCOM Europa.
Anderson, Gregory and K. David Harrison. (to appear) "Natural" and obsolescent change in Tofa, in *Historical, Areal and Typological Aspects of South-Siberian Turkic*, ed. by I. Nevskaya, Wiesbaden: Harrassowitz.
Anderson, J. L., J. L. Morgan, and K. S. White. (2003) A statistical basis for speech sound discrimination. *Language and Speech* 46, 155–182.
Anderson, John M. (1987) Structural analogy and dependency phonology, in *Explorations in Dependency Phonology*, ed. by John M. Anderson and Jacques Durand, 15–47, Dordrecht: Foris.
Anderson, John M. (1986) Suprasegmental dependencies, in *Dependency and Non-linear Phonology*, ed. by Jacques Durand, 55–133, London: Croom Helm.
Anderson, John M. (1992) *Linguistic Representation: Structural, Analogy and Stratification*. Berlin: Mouton de Gruyter.
Anderson, John M. (2004) Contrast in phonology, structural analogy, and the interfaces. *Studia Linguistica*, 58, 269–287.
Anderson, John M. (2006) Structural analogy and universal grammar. *Lingua* 116, 601–633.
Anderson, John M. and Jacques Durand. (1986) Dependency phonology, in *Dependency and Non-linear Phonology*, ed. by Jacques Durand, 1–54, London: Croom Helm.
Anderson, John M. and Jacques Durand (eds.) (1987a) *Explorations in Dependency Phonology*. Dordrecht: Foris.
Anderson, John M. and Jacques Durand. (1987b) Introduction, in *Explorations in Dependency Phonology*, ed. by John M. Anderson and Jacques Durand, 1–13, Dordrecht: Foris.
Anderson, John M. and Colin J. Ewen. (1980a) Introduction: A sketch of dependency phonology, in *Studies in Dependency Phonology*, ed. by John M. Anderson and Colin J. Ewen, 9–40, Ludwigsburg: Strauch.
Anderson, John M. and Colin J. Ewen (eds.) (1980b) *Studies in Dependency Phonology*. Ludwigsburg: R. O. U. Strauch.
Anderson, John M. and Colin J. Ewen. (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, John M., Colin J. Ewen, and Jørgen Staun. (1985) Phonological structure: Segmental, suprasegmental and extrasegmental. *Phonology Yearbook* 2, 203–224.

Anderson, John M. and C. Jones. (1974) Three theses concerning phonological representations. *Journal of Linguistics* 10, 1–26.

Anderson, Stephen R. (1969) *West Scandinavian Vowel Systems and the Ordering of Phonological Rules*. MIT dissertation.

Anderson, Stephen R. (1974) *The Organization of Phonology*. New York: Academic Press.

Anderson, Stephen R. (1981) Why phonology isn't natural. *Linguistic Inquiry* 12, 493–539.

Anderson, Stephen R. (1982) The analysis of French shwa, or how to get something from nothing. *Language* 58, 121–138.

Anderson, Stephen R. (1984) Kwakw'ala syntax and the government/binding theory, in *The Syntax of Native American Languages*, ed. by E.-D. Cook and Donna Gerdts, 21–75, New York: Academic Press.

Anderson, Stephen R. (1985) *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago: University of Chicago Press.

Anderson, Stephen R. (1988) Morphological change, in *Linguistics: The Cambridge Survey*, ed. by Frederick J. Newmeyer, volume 1, 324–362, Cambridge: Cambridge University Press.

Anderson, Stephen R. (1992) *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. (2001) On some issues in morphological exponence, in *Yearbook of Morphology*, ed. by Geert Booij and Jaap van Marle, 1–18, Dordrecht: Kluwer.

Anderson, Stephen R. (2005) *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Andruski, Jean E. and Martha Ratcliff. (2000) Phonation types in production of phonological tone: the case of Green Mong. *Journal of the International Phonetics Association* 30, 63–82.

Angluin, Dana. (1982) Inference of reversible languages. *Journal of the Association for Computing Machinery* 29, 741–765.

Anttila, Arto. (1997a) Deriving variation from grammar: A study of Finnish genitives, in *Variation, Change and Phonological Theory*, ed. by Frans Hinskens, Roeland van Hout, and Leo Wetzels, 35–68, Amsterdam: Benjamins.

Anttila, Arto. (1997b) *Variation in Finnish Phonology and Morphology*. Stanford University dissertation.

Anttila, Arto. (2002a) Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20, 1–42.

Anttila, Arto. (2002b) Variation and phonological theory, in *The Handbook of Language Variation and Change*, ed. by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 206–243, Oxford: Blackwell.

Anttila, Arto. (2006) Variation and opacity. *Natural Language and Linguistic Theory* 24, 893–944.

Anttila, Arto. (2007) Variation, typology, universals. Class handout. Retrieved December 27, 2007 from http://www.stanford.edu/antilla/teaching/lsa-2007/variation.html.

Anttila, Arto and Curtin Andrus. (2006) T-Orders. Manuscript, Stanford University.

Anttila, Arto and Vivienne Fong. (2000) The partitive constraint in Optimality Theory. *Journal of Semantics* 17, 281–314.

Anttila, Arto and Vivienne Fong. (2004) Variation, ambiguity, and noun classes in English. *Lingua* 114, 1253–1290.

Anttila, Arto, Vivienne Fong, Štefan Beňuš, and Jennifer Nycz. (2008) Variation and opacity in Singapore

English consonant clusters. *Phonology* 25, 181–216.
Anttila, Arto and Young-mee Yu Cho. (1998) Variation and change in Optimality Theory. *Lingua* 104, 31–56.
Ao, Benjamin. (1991) Kikongo nasal harmony and context-sensitive underspecification. *Linguistic Inquiry* 22, 193–196.
Aoki, Haruo. (1966) Nez Perce vowel harmony and Proto-Sahaptian vowels. *Language* 42, 759–767.
Aoyama, Katsura. (2003) Perception of syllable-initial and syllable-final nasals in English by Korean and Japanese speakers. *Second Language Learning* 19, 251–265.
Apoussidou, Diana. (2007) *The Learnability of Metrical Phonology*. Utrecht: LOT.
Applegate, Richard. (1972) *Ineseño Chumash Grammar*. University of California, Berkeley dissertation.
Arbisi-Kelm, Timothy and Mary E. Beckman. (2009) Prosodic structure and consonant development across languages, in *Phonetics and Phonology: Interactions and Interrelations*, ed. by Marina Vigário, Sónia Frota, and M. João Freitas, 109–136, Amsterdam: Benjamins.
Archangeli, Diana. (1983) The root CV-template as a property of the affix: Evidence from Yawelmani. *Natural Language and Linguistic Theory* 1, 348–384.
Archangeli, Diana. (1984) *Underspecification in Yawelmani Phonology and Morphology*. MIT dissertation.
Archangeli, Diana. (1988a) Aspects of underspecification theory. *Phonology* 5, 183–208.
Archangeli, Diana. (1988b) *Underspecification in Yawelmani Phonology and Morphology*. New York: Garland Publications.
Archangeli, Diana. (1991) Syllabification and prosodic templates in Yawelmani. *Natural Language and Linguistic Theory* 9, 231–283.
Archangeli, Diana and D. Terence Langendoen. (1997) *Optimality Theory: An Overview*. Oxford: Blackwell.
Archangeli, Diana and Douglas Pulleyblank. (1987) Minimal and maximal rules: Effects of tier scansion, in *Proceedings of the 17th Annual Meeting of the North East Linguistic Society*, 16–35.
Archangeli, Diana and Douglas Pulleyblank. (1989) Yoruba vowel harmony. *Linguistic Inquiry* 20, 173–217.
Archangeli, Diana and Douglas Pulleyblank. (1994) *Grounded Phonology*. Cambridge, MA: MIT Press.
Archangeli, Diana and Douglas Pulleyblank. (2002) Kinande vowel harmony: Domains, grounded conditions and one-sided alignment. *Phonology* 19, 139–188.
Archangeli, Diana and Douglas Pulleyblank. (2007) Harmony, in *Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 772–889, Cambridge: Cambridge University Press.
Archibald, J. (ed.) (1995) *The Acquisition of Non-linear Phonology*. Hillsdale, NJ: Lawrence Erlbaum.
Ard, Joshua. (1981) A sketch of vowel harmony in the Tungus languages, in *Studies in the Languages of the USSR*, ed. by Bernard Comrie, 23–43, Edmonton: Linguistic Research.
Arnason, Kristján. (1980) *Quantity in Historical Phonology: Icelandic and Related Cases*. Cambridge, MA: MIT Press.
Arnott, David W. (1970) *The Nominal and Verbal Systems of Fula*. Oxford: Oxford University Press.
Aronoff, Mark. (1988) Head operations and strata in reduplication: A linear treatment. *Yearbook of Morphology* 1, 1–15.
Aronoff, Mark. (1994) *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
Aronoff, Mark, Irit Meir, Carol Padden, and Wendy Sandler. (2008) The roots

of linguistic organization in a new language. *Interaction Studies* 9, 133–153.
Aronoff, Mark, Irit Meir, and Wendy Sandler. (2005) The paradox of sign language morphology. *Language* 81, 301–344.
Arregi, Karlos. (2002) *Focus on Basque Movements*. MIT dissertation.
Arsenault, Paul and Alexei Kochetov. (to appear) Retroflex harmony in Kalasha: Agreement or spreading?, in *Proceedings of the 39th Annual Meeting of the North East Linguistic Society.*
Arvaniti, Amalia. (1994) Acoustic features of Greek rhythmic structure. *Journal of Phonetics* 22, 239–268.
Arvaniti, Amalia. (1999) Standard Modern Greek. *Journal of International Phonetic Association* 19, 167–172.
Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen. (1998) Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26, 3–25.
Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen. (2000) What is a starred tone? Evidence from Greek, in *Papers in Laboratory Phonology V*, ed. by Michael B. Broe and Janet B. Pierrehumbert, 119–131, Cambridge University Press.
Askins, D. and David Perlmutter. (1995) Allomorphy explained through phonological representation: person and number inflection of American Sign Language. Presented at the annual meeting of the German Linguistic Society.
Aslin, Richard N., Jenny R. Saffran, and Elissa L. Newport. (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9, 321–324.
Atal, Bishnu S. and Suzanne L. Hanauer. (1971) Speech analysis and synthesis by linear prediction of the speech wave. *Journal of the Acoustical Society of America* 50, 637–655.
Auer, Peter. (1994) *Einige Argumente gegen die Silbe als Universale Prosodische Hauptkategorie*. Tübingen: Max Niemeyer.
Auger, Julie. (2001) Phonological variation and Optimality Theory: Evidence from word-initial vowel epenthesis in Vimeu Picard. *Language Variation and Change* 13, 253–303.
Aylett, Mark and Alice Turk. (2004) The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47, 31–56.
Baayen, R. Harald (2007) Storage and computation in the mental lexicon, in *The Mental Lexicon: Core Perspectives*, ed. by Gonia Jarema and Gary Libben, 81–104, Oxford: Elsevier.
Baayen, R. Harald. (2005a) Corpus linguistics in morphology: Morphological productivity, in *Handbook of Corpus Linguistics*, ed. by A. Lüdeling, M. Kytö, and T. McEnery, De Gruyter.
Baayen, R. Harald. (2005b) Data mining at the intersection of psychology and linguistics, in *Twenty-first Century Psycholinguistics: Four Cornerstones*, ed. by Anne Cutler, 69–83, Hillsdale, NJ: Erlbaum.
Baayen, R. Harald. (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
Baayen, R. Harald, Tom Dijkstra, and Rob Schreuder. (1997) Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language* 36, 94–117.
Baayen, R. Harald, W. M. J. Levelt, R. Schreuder and Mirjam Ernestus. (2008) Paradigmatic structure in speech production, in *Papers from the 43rd Annual Meeting of the Chicago Linguistic Society*, ed. by Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki, and Suwon Yoon, volume 1, 1–29, Chicago: Chicago Linguistic Society.
Baayen, R. Harald, Richard Piepenbrock, and Leon Gulikers. (1995) *The CELEX*

Database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium.
Baayen, R. Harald, Lee H. Wurm, and Joanna Aycock. (2007) Lexical dynamics for low-frequency complex words. A regression study across tasks and modalities. *The Mental Lexicon* 2, 419–463.
Bach, Emmon. (1968) Two proposals concerning the simplicity metric in phonology. *Glossa* 2, 127–149.
Bach, Emmon and Robert T. Harms. (1972) How do languages get crazy rules?, in *Linguistic Change and Generative Theory*, ed. by Robert Stockwell and Ronald K. S. Macaulay, 1–21, Bloomington, IN: Indiana University Press.
Bach, Emmon and W. Marsh. (1987) An elementary proof of the Peters-Ritchie theorem, in *The Formal Complexity of Natural Language*, ed. by W. J. Savitch, Emmon Bach, W. Marsh, and G. Safran-Naveh, 41–55, Dordrecht: Kluwer.
Bach, Emmon and Deirdre Wheeler. (1981) Montague phonology: A first approximation, in *University of Massachusetts Occasional Papers in Linguistics, VII* 27–45, University of Massachusetts.
Baddeley, Alan. (2003) Working memory: Looking back and looking forward. *Nature Reviews Neuroscience* 4, 829–839.
Baerman, Matthew. (2005) Directionality and (Un)natural Classes in Syncretism. *Language* 80, 807–827.
Baertsch, Karen. (2002) *An Optimality Theoretic Approach to Syllable Structure: The Split Margin Hierarchy*. Indiana University dissertation.
Bagemihl, Bruce. (1988) *Alternative Phonologies and Morphologies*. University of British Columbia dissertation.
Bagemihl, Bruce. (1991) Syllable structure in Bella Coola. *Linguistic Inquiry* 22, 589–646.
Bagemihl, Bruce. (1995) Language games and related areas, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 697–712, Oxford: Blackwell.
Bahan, Benjamin. (1996) *Nonmanual Realization of Agreement in American Sign Language*. Boston University dissertation.
Bailey, Todd M. and Ulrike Hahn. (2001) Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44, 568–591.
Bakalla, Mohammed. (1973) *The Morphology and Phonology of Meccan Arabic*. School of Oriental and African Studies, University of London dissertation.
Baker, Brett. (1997) Edge crispness: Segment to mora isomorphism, in *Proceedings of the 16th Annual Meeting of the West Coast Conference on Formal Linguistics*, 33–47.
Baker, Brett. (2008) *Word Structure in Ngalakgan*. Stanford: CSLI.
Baker, William and Bruce L. Derwing. (1982) Response coincidence analysis as evidence for language acquisition strategies. *Applied Psycholinguistics* 3, 193–221.
Baković, Eric. (1996) Foot harmony and quantitative adjustments. Manuscript, Rutgers University.
Baković, Eric. (1998) Unbounded stress and factorial typology. Rutgers Optimality Archive 244.
Baković, Eric. (2000) *Harmony, Dominance and Control*. Rutgers University dissertation. ROA 360.
Baković, Eric. (2002) Vowel harmony and cyclicity in Eastern Nilotic, in *Proceedings of the 27th Annual Meeting of the Berkeley Linguistics Society.*
Baković, Eric. (2003) Vowel harmony and stem identity, in *San Diego Linguistic Papers*, volume 1, 1–42.
Baković, Eric. (2005) Antigemination, assimilation and the determination of identity. *Phonology* 22, 279–315.
Baković, Eric. (2007) A revised typology of opaque generalisations. *Phonology* 24, 217–259.

Baković, Eric. (forthcoming) *On "Elsewhere": Disjunctivity and Blocking in Phonological Theory*. London: Equinox Publishing.

Baković, Eric and Bożena Pająk. (2008) Why and how not to counterbleed. Presented at the *Workshop on Phonological Voicing Variation*, Amsterdam/Leiden. Available online.

Baković, Eric and Colin Wilson. (2001) Transparency, strict locality and targeted constraints, in *Proceedings of the 19th West Coast Conference on Formal Linguistics*, 43–56.

Ball, Martin J. (1997) Monovalent phonologies: Dependency phonology and an introduction to government phonology, in *The New Phonologies: Directions in Clinical Linguistics*, ed. by M. J. Ball and R. D. Kent, 127–161, San Diego, CA: Singular Publishing.

Balling, Laura and R. Harald Baayen. (2008) Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes* 23, 1159–1190.

Balog, Heather and Diane Brentari. (2008) The relationship between early gestures and intonation. *First Language* 18, 141–163.

Baltaxe, Christiane. (1969) *Principles of Phonology*. Berkeley: University of California Press. Translation of Trubetzkoy (1939).

Banfield, A. W. (1914) *Dictionary of the Nupe language*. The Niger Press.

Bao, Zhiming. (1999) *The Structure of Tone*. Oxford: Oxford University Press.

Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. (1996) Magnitude estimation of linguistic acceptability. *Language* 72, 32–68.

Barker, M. A. R. (1964) *Klamath Grammar*. Berkeley, CA: University of California Press.

Barlow, Jessica. (1997) *A Constraint-Based Account of Syllable Onsets: Evidence from Developing Systems*. Indiana University dissertation.

Barlow, Jessica. (2001a) Individual differences in the production of initial consonant sequences in Pig Latin. *Lingua* 111, 667–696.

Barlow, Jessica. (2001b) The structure of /s/-sequences: evidence from a disordered system. *Journal of Child Language* 28, 291–324.

Barlow, Jessica. (2007) Grandfather effects: A longitudinal case study of the phonological acquisition of intervocalic consonants in English. *Language Acquisition* 14, 121–164.

Barnes, Janet. (1996) Autosegments with three-way contrasts in Tuyuca. *International Journal of American Linguistics* 62, 31–58.

Barnes, Jonathan. (2006) *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin: Mouton de Gruyter.

Barnes, Jonathan and Darya Kavitskaya. (2003) Phonetic analogy and schwa deletion in French, in *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, ed. by Julie Larsen and Mary Paster, 39–50, Berkeley.

Barnes, Jonathan. (2002) *Positional Neutralization: A Phonologization Approach to Typological Patterns*. Berkeley: University of California, Berkeley dissertation.

Baroni, Marco. (2000) *Distributional Cues in Morpheme Discovery: A Computational Model and Empirical Evidence*. UCLA dissertation.

Baroni, Marco and Laura Vanelli. (2000) The relations between vowel length and consonantal voicing in Friulian, in *Phonological Theory and the Dialects of Italy*, ed. by Lori Repetti, Amsterdam and Philadelphia: Benjamins.

Barr, Robin. (1994) *A Lexical Model of Morphological Change*. Harvard University dissertation.

Barry, Martin. (1985) A palatographic study of connected speech processes. Technical report. Cambridge Papers in Phonetics and Experimental Linguistics.

Basbøll, Hans. (2005) *The Phonology of Danish*. Oxford: Oxford University Press.

Battison, Robbin. (1978) *Lexical Borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.

Baudouin de Courtenay, Jan. (1895) *Versuch einer Theorie phonetischer Alternationen. Ein Capitel aus der Psychophonetik*. Strassburg: Trübner. [English translation ("An attempt at a theory of phonetic alternations") published in *A Baudouin de Courtenay Anthology: The Beginnings of Structural Linguistics*, ed. by Edward Stankiewicz, 144–212. Bloomington, IN: Indiana University Press, 1972].

Bauer, Laurie. (2003) *Introducing Linguistic Morphology*. Washington, DC: Georgetown University Press, 2nd edition.

Bauer, Robert S. and Paul K. Benedict. (1997) *Modern Cantonese Phonology*. Berlin and New York: Mouton de Gruyter.

Bayley, Robert. (1995) Consonant cluster reduction in Tejano English. *Language Variation and Change* 6, 303–326.

Bayley, Robert. (1997) Variation in Tejano English: Evidence for variable lexical phonology, in *Language Variety in the South Revisited*, ed. by C. Bernstein, T. Nunnaly, and R. Sabino, 197–209, Tuscaloosa, AL: University of Alabama Press.

Bayley, Robert. (2002) The quantitative paradigm, in *The Handbook of Language Variation and Change*, ed. by J. K. Chambers, Peter Trudgill, Nathalie Schilling-Estes, 117–141, Oxford: Blackwell.

Bear, John. (1990) Backwards phonology, in *COLING-90: Papers presented to the 13th International Conference on Computational Linguistics*, ed. by Hans Karlgren, volume III, 13–20.

Beaulieu, Kathy. (2004) *L'adaptation des emprunts anglais en FQ de Québec: Analyse de la graphie en contexte hors-malformation*. Quebec City: Project CoPho Research Report, Laval University.

Becker, Michael. (2009) *Phonological Trends in the Lexicon*. University of Massachusetts, Amherst dissertation.

Becker, Michael, Nihan Ketrez, and Andrew Nevins. (2011) The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish voicing deneutralization. *Language* 87, 84–125.

Becker, Michael, Joe Pater and Christopher Potts. (2007) OT-Help 1.1. Retrieved April 14, 2011 from http://web.linguist.umass.edu/OTHelp/.

Beckman, Jill. (1995) Shona height harmony: Markedness and positional identity, in *University of Massachusetts Occasional Papers: Papers in Optimality Theory 18*, ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 53–75, Amherst, MA: GLSA.

Beckman, Jill. (1997) Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14, 1–46.

Beckman, Jill. (1998) *Positional Faithfulness*. University of Massachusetts, Amherst dissertation. Published by Garland Publishing, New York, 1999.

Beckman, Mary E. (1986) *Stress and Non-Stress Accent*. Dordrecht: Foris.

Beckman, Mary E., Donna Erickson, Kiyoshi Honda, Hiroyuki Hirai, and Seiji Niimi. (1995) Physiological correlates of global and local pitch range variation in the production of high tones in English, in *Proceedings of the 13th International Congress of Phonetics Sciences*, volume 2, 638–641.

Beckman, Mary E., Susan R. Hertz and Osamu Fujimura. (1983) SRS pitch rules for Japanese. *Working Papers of the Cornell Phonetics Laboratory* 1, 1–16.

Beckman, Mary E., Julia Hirschberg, and Stefanie Shattuck-Hufnagel. (2005) The original ToBI system and the evolution of the ToBI framework, in *Prosodic Typology: The Phonology of Intonation and Phrasing*, ed. by Sun-Ah Jun, 9–54, Oxford: Oxford University Press.

Beckman, Mary E. and John Kingston. (1990) Introduction, in *Papers in*

Laboratory Phonology I: Between the Grammar and Physics of Speech, ed. by John Kingston and Mary E. Beckman, 1–16, Cambridge: Cambridge University Press.

Beckman, Mary E. and Janet B. Pierrehumbert. (1986) Intonational structure in Japanese and English. *Phonology Yearbook* 3, 225–309.

Beckman, Mary E. and Janet B. Pierrehumbert. (1992) Strategies and tactics for thinking about F0 variation, in *Papers in Laboratory Phonology II*, ed. by Gerard J. Docherty and D. Robert Ladd, 387–397, Cambridge: Cambridge University Press.

Beddor, Patrice S., James D. Harnsberger, and Stephanie Lindemann. (2002) Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics* 30, 591–627.

Beddor, Patrice S. and Rena A. Krakow. (1999) Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *The Journal of the Acoustical Society of America* 106, 2868–2887.

Beddor, Patrice S., Rena A. Krakow, and Louis M. Goldstein. (1986) Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology Yearbook* 3, 197–217.

Beddor, Patrice S., Rena A. Krakow, and Stephanie Lindemann. (2001) Patterns of perceptual compensation and their phonological consequences, in *The Role of Speech Perception in Phonology*, ed. by Elizabeth V. Hume and Keith Johnson, 55–78, San Diego, CA: Academic Press.

Bell, Alan. (1976) Review of *Syllable, Word, Nexus, Cursus* by Ernst Pulgram. *Language* 52(1), 241–245.

Bell, Alan. (1978) Syllabic consonants, in *Universals of Human Language*, ed. by Joseph Greenberg, volume 2: *Phonology*, 153–202, Stanford, CA: Stanford University Press.

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, C. Girand, M. Gregory and Daniel Gildea. (2003) Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113, 1001–1024.

Bell, C. G., Hiroya Fujisaki, J. M. Heinz, Kenneth N. Stevens, and Arthur S. House. (1961) Reduction of speech spectra by analysis-by-synthesis techniques. *Journal of the Acoustical Society of America* 33, 1725–1736.

Benedict, Paul K. (1948) Tonal systems in Southeast Asia. *Journal of the American Oriental Society* 64, 184–191.

Benedicto, Elena and Diane Brentari. (2004) Where did all the arguments go? Argument-changing properties of classified in ASL. *Natural Language and Linguistic Theory* 22, 743–810.

Benua, Laura. (1995) Identity effects in morphological truncation, in *Papers in Optimality Theory*, ed. by Jill Beckman, Suzanne Urbanczyk, and L. Walsh, Amherst, MA: Graduate Linguistic Student Association.

Benua, Laura. (2000) *Phonological Relations Between Words*. Abingdon, UK: Routledge.

Benus, Stefan and Adamantios Gafos. (2007) Articulatory characteristics of Hungarian "transparent" vowels. *Journal of Phonetics* 35, 271–300.

Benus, Stefan, Adamantios Gafos and Louis Goldstein. (2004) Phonetics and phonology of transparent vowels in Hungarian, in *Proceedings of the 29th Annual Meeting of the Berkeley Linguistics Society*, 485–497.

Berent, Iris, Daniel L. Everett, and Joseph Shimron. (2001) Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology* 42, 1–60.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel A. Moreno, and Paul Smolensky. (2008) Language universals in human

brains, in *Proceedings of the National Academy of Sciences*, volume 105, 5321–5325.
Berent, Iris, Tracy Lennertz, Paul Smolensky, and Vered Vaknin-Nusbaum. (2009) Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26, 75–108.
Berent, Iris and Joseph Shimron. (1997) The representation of Hebrew words: Evidence from the Obligatory Contour Principle. *Cognition* 64, 39–72.
Berent, Iris and Joseph Shimron. (2003) Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics* 39, 31–55.
Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. (2007) What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104, 591–630.
Berg, Rob van den, Carlos Gussenhoven, and Toni Rietveld. (1992) Downstep in Dutch: Implications for a model, in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, ed. by Gerard J. Docherty and D. Robert Ladd, 335–367, Cambridge University Press.
Berko, Jean. (1958) The child's acquisition of English morphology. *Word* 14, 150–177.
Bermúdez-Otero, Ricardo. (1999) *Constraint Interaction in Language Change: Quantity in English and Germanic*. University of Manchester and Universidad de Santiago de Compostela dissertation.
Bermúdez-Otero, Ricardo. (2003) The acquisition of phonological opacity, in *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, ed. by J. Spenader, A. Eriksson, and O. Dahl, 25–36, Stockholm, Department of Linguistics, Stockholm University.
Bermúdez-Otero, Ricardo. (to appear) *Stratal Optimality Theory*. Oxford: Oxford University Press.
Bermúdez-Otero, Ricardo and P. Honeybone. (2006) Special issue: *Phonology and Syntax – The Same or Different? Lingua* 116.
Bernhardt, B. H., B. Gick, P. Bacsfalvi, and M. Adler-Bock. (2005) Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics and Phonetics* 19, 605–617.
Bernhardt, B. H., J. Gilbert, and D. Ingram. (eds.) (1996) *Proceedings of the UBC International Conference on Phonological Acquisition*, Somerville, MA: Cascadilla Press.
Bernhardt, B. H. and J. P. Stemberger. (1998) *Handbook of Phonological Development: From the Perspective of Constraint-based Nonlinear Phonology*. San Diego: Academic Press.
Bertinetto, Pier Marco. (1992) The use and misuse of external evidence in phonology, in *Phonologica 1988*, ed. by Wolfgang U. Dressler, Hans Luschützky, Oskar Pfeiffer, and John Rennison, 33–47, Cambridge University Press, New York.
Bertinetto, Pier Marco. (2004) On the undecidable syllabification of /sc/ clusters in Italian: Converging experimental evidence. *Rivista di Linguistica* 16, 349–372.
Bertram, Raymond M., Matti Laine, R. Harald Baayen, R. Schreuder, and Jukka Hyönä. (1999) Affixal homonymy triggers full-form storage, even with inflected words, even in a morphologically rich language. *Cognition* 74, B13–B25.
Berwick, Robert C. (1986) Learning from positive-only examples: The subset principle and three case studies, in *Machine Learning: An Artificial Intelligence Approach: Volume II*, ed. by R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, 625–645, Los Altos, CA: Kaufmann.
Bessell, Nicola. (1998) Local and non-local consonant-vowel interaction in Interior Salish. *Phonology* 15, 1–40.
Best, Catherina T., G. W. McRoberts, and N. M. Sithole. (1988) Examination of

perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 345–360.

Best, Catherina T., Gerald W. McRoberts, and Elizabeth Goodell. (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America* 109, 775–794.

Best, Catherine T. and Winifred Strange. (1992) Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20, 305–330.

Bethin, Christina. (1978) *Phonological Rules in the Nominative Singular and Genitive Plural of the Slavic Substantive Declension*. University of Illinois Champaign-Urbana dissertation.

Bickel, Balthasar and Johanna Nichols. (2007) Inflectional morphology, in *Language Typology and Syntactic Description*, ed. by Timothy Shopen, Cambridge: Cambridge University Press.

Bickmore, Lee. (2007) Branching nodes and prosodic categories: Evidence from Kinyambo, in *The Phonology-Syntax Connection*, ed. by S. Inkelas and D. Zec, Chicago: University of Chicago Press.

Birch, Barbara. (2002) *English L2 Reading: Getting to the Bottom*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bird, Steven. (1995) *Computational Phonology: A Constraint-based Approach*. Cambridge: Cambridge University Press.

Bird, Steven. (1998) Strategies for representing tone in African writing systems: A critical review. *Written Language and Literacy* 2, 1–44.

Bird, Steven. (1999) When marking tone reduces fluency: An orthography experiment in Cameroon. *Language and Speech* 42, 83–115.

Bird, Steven, John Coleman, Janet Pierrehumbert, and James Scobbie. (1992) Declarative Phonology, in *Proceedings of the 15th International Conference of Linguists*, Quebec, Canada.

Bird, Steven and T. Mark Ellison. (1994) One-level phonology: Autosegmental representations and rules as finite state automata. *Computational Linguistics* 20, 55–90.

Bird, Steven and Ewen Klein. (1994) Phonological analysis in typed feature systems. *Computational Linguistics* 20, 455–491.

Bishop, Jason. (2007) Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differences involved in *s*-aspiration, in *Proceedings of the International Congress of Phonetic Sciences XVI*, 1765–1768.

Blevins, Juliette. (1995) The syllable in phonological theory, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 206–244, Oxford: Blackwell.

Blevins, Juliette. (1997) Rules in Optimality Theory: Two case studies, in *Derivations and Constraints in Phonology*, ed. by Iggy Roca, 227–260, Oxford: Clarendon Press.

Blevins, Juliette. (1999) Untangling Leti infixation. *Oceanic Linguistics* 38, 383–403.

Blevins, Juliette. (2003) The independent nature of phonotactic constraints: An alternative to syllable-based approaches, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, 375–403, Cambridge: Cambridge University Press.

Blevins, Juliette. (2004) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette. (2005) Understanding antigemination: Natural or unnatural history, in *Linguistic Diversity and Language Theories*, ed. by Zygmunt

Frajzyngier, 203–234, Amsterdam: Benjamins.
Blevins, Juliette. (2006a) Reply to commentaries. *Theoretical Linguistics* 32, 245–256.
Blevins, Juliette. (2006b) Syllable typology, in *Encyclopedia of Language and Linguistics*, volume 12, 333–337, Elsevier.
Blevins, Juliette. (2006c) A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32, 117–166.
Blevins, Juliette and Andrew Garrett. (1998) The origins of consonant-vowel metathesis. *Language* 74, 508–556.
Blevins, Juliette and Andrew Garrett. (2004) The evolution of metathesis, in *Phonetically-Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 117–156, Cambridge University Press.
Bloch, Bernard. (1941) Phonemic overlapping. *American Speech* 16, 278–284.
Bloch, Bernard. (1948) A set of postulates for phonemic analysis. *Language* 24, 3–46.
Bloomfield, Leonard. (1933) *Language*. New York: H. Holt.
Blumenfeld, Lev. (2006) *Constraints on Phonological Interactions*. Stanford University dissertation.
Blumstein, Sheila. (1973) *A Phonological Investigation of Aphasic Speech*. The Hague: Mouton.
Blust, Robert. (1996) Low vowel dissimilation in Ere. *Oceanic Linguistics* 36, 96–112.
Boas, Franz. (1947) Grammar of Kwakiutl (with a glossary of the suffixes). *Transactions of the American Philosophical Society* 37, 201–377.
Bobaljik, Jonathan. (2000) The ins and outs of contextual allomorphy, in *University of Maryland Working Papers in Linguistics*, ed. by K. K. Grohmann and C. Struijke, volume 10, 35–71, Department of Linguistics, University of Maryland.
Bochner, Harry. (1993) *Simplicity in Generative Morphology*. Berlin: Mouton de Gruyter.
Bod, Rens, Jennifer Hay, and Stefanie Jannedy. (2003) *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
Bodomo, Adams. (2006) The structure of ideophones in African and Asian Languages: The case of Dagaare and Cantonese, in *Selected Proceedings of the 35th Annual Conference on African Languages*, ed. by J. Mugane, 203–213, Somerville, MA: Cascadilla Press.
Boersma, Paul. (1993) Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound, in *Proceedings of the Institute of Phonetic Sciences, Amsterdam*, volume 17, 97–110.
Boersma, Paul. (1997) How we learn variation, optionality, and probability, in *Institute of Phonetic Sciences, University of Amsterdam, Proceedings* 21, 43–58.
Boersma, Paul. (1998) *Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives*. University of Amsterdam dissertation. The Hague: Holland Academic Graphics.
Boersma, Paul. (1999) On the need for a separate perception grammar. ROA 358.
Boersma, Paul. (2003) Nasal harmony in functional phonology, in *The Phonological Spectrum*, ed. by Jeroen van de Weijer, Vincent J. van Heuven, and Harry van der Hulst, volume 1, 3–35, Amsterdam: Benjamins.
Boersma, Paul and Silke Hamann. (2008) The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25, 217–270.
Boersma, Paul and Bruce Hayes. (2001) Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32, 45–86.
Boersma, Paul and C. Levelt. (2000) Gradual constraint-ranking learning algorithm predicts acquisition order, in *Proceedings of Child Language Research Forum*, volume 30, 229–237, Stanford, CA: CSLI Publications.

Boersma, Paul and Joe Pater. (2008) Convergence properties of a gradual learner in harmonic grammar. ROA 970.

Boersma, Paul and David Weenink. (2007) Praat: Doing Phonetics by Computer (Version 5.0.01). [Computer program]. Retrieved December 18, 2007, from http://www.praat.org/.

Bolinger, Dwight. (1958) A theory of pitch accent in English. *Word* 14, 109–149.

Bolinger, Dwight. (1978) Intonation across languages, in *Universals of Human Language, Vol. 2: Phonology*, ed. by Joseph Greenberg, 471–524, Stanford, CA: Stanford University Press.

Bonatti, Luca, Marcela Peña, Marina Nespor, and Jacques Mehler. (2005) Linguistic constraints on statistical computations. *Psychological Science* 16, 451–459.

Bonte, Milene L., Holger Mitterer, Nadia Zellagui, Hane Poelmans and Leo Blomert. (2005) Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology* 116, 2765–2774.

Bonvillian, John D., Michael D. Orlansky and Raymond J. Folven. (1990) Early sign language acquisition: Implications for theories of language acquisition, in *From Gesture to Language in Hearing and Deaf Children*, ed. by Virginia Volterra and Carol J. Erting, 219–232, Berlin: Springer Verlag.

Booij, Geert. (1981) *Generatieve Fonologie van het Nederlands*. Utrecht: Het Spectrum.

Booij, Geert. (1984) Coordination reduction in complex words: A case for prosodic phonology, in *Advances in Non-linear Phonology*, ed. by Harry van der Hulst and Norval Smith, 143–160, Dordrecht: Foris.

Booij, Geert. (1995) *The Phonology of Dutch*. Oxford: Clarendon Press.

Booij, Geert. (1996) Cliticization as prosodic integration: The case of Dutch. *The Linguistic Review* 13, 219–242.

Booij, Geert. (2000a) The balance between storage and computation in phonology in *Storage and Computation in the Language Faculty*, ed. by Sieb Nooteboom, Fred Weerman, and Frank Wijnen, 133–156, Dordrecht: Kluwer Academic Publishers.

Booij, Geert. (2001) Prosodic restrictions on affixation in Dutch, in *Yearbook of Morphology 2001*, ed. by Geert Booij and Jaap van Marle, 183–201, Dordrecht: Kluwer.

Booij, Geert. (2002b) *The Morphology of Dutch*. Oxford: Oxford University Press.

Booij, Geert. (2005) The interface between phonology and morphology: A comment on Ellen Kaisse, "Word formation and phonology." *Skase Journal of Theoretical Linguistics* 2, 17–25.

Booij, Geert and R. Lieber. (1993) On the simultaneity of morphological and prosodic structure, in *Studies in Lexical Phonology*, ed. by Ellen Kaisse and Sharon Hargus, 23–44, New York: Academic Press.

Booker, Karen M. (1980) *Comparative Muskogean: Aspects of Proto-Muskogean Verb Morphology*. University of Kansas dissertation.

Bortfeld, H., J. L. Morgan, R. M. Golinkoff and K. Rathbun. (2005) Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science* 16, 298–304.

Botma, Bert. (2004) *Phonological Aspects of Nasality: An Element-based Dependency Approach*. University of Leiden dissertation. Published by LOT, no. 90.

Botne, Robert and Stuart Davis. (2000) Language games, segment imposition, and the syllable. *Studies in Language* 24, 319–344.

Boucher, V. (1988) A parameter of syllabification for VstopV and relative-timing invariance. *Journal of Phonetics* 16, 299–326.

Boyce, S. E. (1988) *The Influence of Phonological Structure on Articulatory Organization in Turkish and in English: Vowel Harmony and Coarticulation*. Yale University dissertation.

Braconnier, Cassian. (1982) Le système tonal du dioula d'Odienné, tome 1. Technical report, Institut de Linguistique Appliquée.

Bradley, C. Henry. (1970) *A Linguistic Sketch of Jicaltepec Mixtec*. Summer Institute of Linguistics of the University of Oklahoma.

Braem, Penny Boyes. (1981) *Distinctive Features of the Handshapes of American Sign Language*. University of California dissertation.

Braem, Penny Boyes. (1999) Rhythmic and temporal patterns in the signing of deaf early and late signers of Swiss German Sign Language. *Language and Speech* 42, 177–208.

Braem, Penny Boyes. (2001) Functions of the mouthing component in the signing of Deaf and early and late learners of Swiss German Sign Language, in *Foreign Vocabulary in Sign Languages*, ed. by Diane Brentari, 1–47, Mahwah, NJ: Lawrence Erlbaum Associates.

Brazell, Charles Ernst, J. C. Catford, M. A. K. Halliday and R. H. Robins (eds.) (1966) *In Memory of J. R. Firth.* London: Longmans, Green, and Co.

Breeze, Mary. (1990) A sketch of the phonology and grammar of Gimira (Benchnon), in *Omotic Language Studies*, ed. by Richard Hayward, 1–67, London: School of Oriental and African Studies.

Bregman, Albert S. (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Breiman, Leo, Jerome Friedman, Charles J. Stone and R. A. Olshen. (1984) *Classification and Regression Trees*. Monterey, CA: Chapman and Hall.

Brennan, Mary. (1990) *Word Formation in British Sign Language*. University of Stockholm dissertation.

Brennan, Mary. (2005) Conjoining word and image in British Sign Language (BSL): An exploration of metaphorical signs in BSL. *Sign Language Studies* 9, 360–382.

Brent, Michael and Timothy Cartwright. (1996) Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61, 93–125.

Brentari, Diane. (1988) Backwards verbs in ASL: Agreement re-opened, in *Papers from the 24th Annual Meeting of the Chicago Linguistic Society: Parasession on Agreement in Grammatical Theory*, ed. by L. MacLeod, G. Larson, and D. Brentari, 16–27, Chicago: University of Chicago.

Brentari, Diane. (1990a) Licensing in ASL handshape change, in *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, 57–68, Washington, DC: Gallaudet University Press.

Brentari, Diane. (1990b) *Theoretical Foundations of American Sign Language Phonology*. University of Chicago dissertation.

Brentari, Diane. (1993) Establishing a sonority hierarchy in American Sign Language: The use of simultaneous structure in phonology. *Phonology* 10, 281–306.

Brentari, Diane. (1994) Prosodic constraints in American Sign Language, in *Sign Language Research*, ed. by H. Bos and T. Schermer, 39–51, Hamburg: Signum Press.

Brentari, Diane. (1995) Sign language phonology: ASL, in *Handbook of Phonological Theory*, ed. by John A. Goldsmith, 615–639, Oxford: Blackwell.

Brentari, Diane. (1998) *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.

Brentari, Diane. (2002) Modality differences in sign language phonology and morphophonemics, in *Modality in Language and Linguistic Theory*, ed. by Richard Meier, D. Quinto-Pozos and K. Cormier, 35–64, Cambridge: Cambridge University Press.

Brentari, Diane. (2005) The use of morphological templates to specify handshapes in sign languages. *Linguistische Berichte* 13, 145–177.

Brentari, Diane. (2006) Effects of language modality on word segmentation: An experimental study of phonological factors in a sign language, in *Papers in Laboratory Phonology VIII*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 155–164, The Hague: Mouton de Gruyter.

Brentari, Diane. (2007) Sign language phonology: Issues of iconicity and universality, in *Verbal and Signed Languages, Comparing Structures, Constructs and Methodologies*, ed. by E. Pizzuto, P. Pietrandrea and R. Simone, 59–80, Berlin: Mouton de Gruyter.

Brentari, Diane. (2010) *Sign Languages: A Cambridge Language Survey*. Cambridge: Cambridge University Press.

Brentari, Diane and Laurinda Crossley. (2002) Prosody on the hands and face: Evidence from American Sign Language. *Sign Language and Linguistics* 5, 105–130.

Brentari, Diane, Carolina González, Amanda Seidl, and Ronnie Wilbur. (2011) Sensitivity to visual prosodic cues in signers and nonsigners. *Language and Speech* 54, 49–72.

Bresnan, Joan. (1982) *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Brockhaus, Wiebke. (1995) *Final Devoicing in the Phonology of German*. Tübingen: Niemeyer.

Bromberger, Sylvain and Morris Halle. (1989) Why phonology is different. *Linguistic Inquiry* 20, 51–70.

Bromberger, Sylvain and Morris Halle. (1997) The contents of phonological signs: A comparison between their use in derivational theories and in optimality theories, in *Derivations and Constraints in Phonology*, ed. by Iggy Roca, 93–123, Oxford: Clarendon Press.

Broselow, Ellen. (1976) *The Phonology of Egyptian Arabic*. University of Massachusetts, Amherst dissertation.

Broselow, Ellen. (1983) Non-obvious transfer: On predicting epenthesis errors, in *Language Transfer in Language Learning*, ed. by Susan M. Gass and Larry Selinker, 269–280, Rowley, MA: Newbury House.

Broselow, Ellen. (1992) Parametric variation in Arabic dialect phonology, in *Perspectives on Arabic Linguistics IV*, ed. by Ellen Broselow, Mushira Eid, and John J. McCarthy, 7–45, Amsterdam and Philadelphia: Benjamins.

Broselow, Ellen. (1995) Skeletal positions and moras, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 175–205, Oxford: Blackwell.

Broselow, Ellen, Su-I Chen, and Marie Huffman. (1997) Syllable weight: Convergence of phonology and phonetics. *Phonology* 14, 47–82.

Broselow, Ellen and John J. McCarthy. (1983) A theory of internal reduplication. *The Linguistic Review* 3, 25–88.

Browman, Catherine P. and Louis M. Goldstein. (1986) Towards an articulatory phonology, in *Phonology Yearbook*, volume 3, 219–252, Cambridge: Cambridge University Press.

Browman, Catherine P. and Louis M. Goldstein. (1988) Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140–155.

Browman, Catherine P. and Louis M. Goldstein. (1989) Articulatory gestures as phonological units. *Phonology* 6, 201–251.

Browman, Catherine P. and Louis M. Goldstein. (1990) Tiers in articulatory phonology, with some implications for casual speech, in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, ed. by John Kingston and Mary E. Beckman, 341–376, Cambridge: Cambridge University Press.

Browman, Catherine P. and Louis M. Goldstein. (1992) Articulatory phonology: An overview. *Phonetica* 49, 155–180.

Brown, Heather, N. K. Sharma, and Kim Kirsner. (1984) The role of script

and phonology in lexical representation. *Quarterly Journal of Experimental Psychology* 36A, 491–505.
Brown, Roger W. and Donald C. Hildum. (1956) Expectancy and the perception of syllables. *Language* 32, 411–419.
Bruce, Gösta. (1977) *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.
Bruce, Gösta. (1982) Developing the Swedish intonation model. *Working Papers, Department of Linguistics, Lund University* 23, 51–116.
Bruce, Gösta. (1983) Accentuation and timing in Swedish. *Folia Linguistica* 17, 221–238.
Bruce, Gösta. (1987) How floating is focal accent?, in *Nordic Prosody, IV*, ed. by Kirsten Gregersen and Hans Basbøll, 41–49, Odense: Odense University Press.
Bruce, Gösta. (1989) Report from the IPA working group on suprasegmental categories. Technical report, Lund University, Department of Linguistics.
Bruce, Gösta. (1990) Alignment and composition of tonal accents: Comments on Silverman and Pierrehumbert's paper, in *Papers in Laboratory Phonology I*, 107–114, Cambridge: Cambridge University Press.
Bruening, Benjamin. (1997) Abkhaz-Mabkhaz: M-reduplication in Abkhaz, weightless syllables, and base-reduplicant correspondence, in *PF: Papers at the Interface*, ed. by M. McGinnis, B. Bruening, and Y. Kang, *MIT Working Papers in Linguistics*, volume 30, Cambridge MA.
Buchwald, Adam B. (2009) Minimizing and optimizing structure in phonology: Evidence from aphasia. *Lingua* 119, 1380–1395.
Buchwald, Adam B., Brenda Rapp, and Maureen Stone. (2007) Insertion of discrete phonological units: An articulatory and acoustic investigation of aphasic speech. *Language and Cognitive Processes* 22, 910–948.
Buckley, Eugene L. (1994) *Theoretical Aspects of Kashaya Phonology and Morphology*. Stanford: CSLI Publications.
Buckley, Eugene L. (2000) On the naturalness of unnatural rules, in *Proceedings from the Second Workshop on American Indigenous Languages*, volume 9, University of California, Santa Barbara Working Papers in Linguistics.
Buckley, Eugene L. (2001) Polish o-raising and phonological explanation. Presented at LSA meeting.
Bueno-Holle, Juan. (2009) Isthmus Zapotec tone and stress. Master's thesis, University of Chicago.
Burton, Martha W. and Karen E. Robblee. (1997) A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25, 97–114.
Burzio, Luigi. (1996) Surface constraints versus underlying representation, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks, 97–122, Paris and University of Salford: University of Salford Publications.
Burzio, Luigi. (1997) Cycles, non-derived environment blocking, and correspondence, in *The Pointing Finger: Conceptual Studies in Optimality Theory*, ed. by J. Dekkers, F. van der Leeuw, and J. van de Weijer, Oxford: Oxford University Press.
Burzio, Luigi. (2002) Segmental contrast meets output-to-output faithfulness, in *Resolving Conflicts in Grammars: Optimality Theory in Syntax, Morphology, and Phonology*, ed. by G. Fanselow and C. Féry, 163–182, Hamburg: Buske Verlag.
Burzio, Luigi. (forthcoming) Derived environment effects, in *The Blackwell Companion to Phonology*, ed. by Marc van Oostendorp, Colin Ewen, Elizabeth V. Hume and Keren Rice, Oxford: Wiley-Blackwell.
Butskhrikidze, Marika. (2002) *The Consonant Phonotactics of Georgian*.

University of Leiden dissertation. Published by LOT, no. 63.
Butt, John and Carmen Benjamin. (2008) *A New Reference Grammar of Modern Spanish*. Chicago: NTC Publishing Group, 3rd edition.
Butt, M. (1992) Sonority and the explanation of syllable structure. *Linguistische Berichte* 137, 45–67.
Butterworth, Brian and Steven Whittaker. (1980) Peggy Babcock's relatives, in *Tutorials in Motor Behavior*, ed. by George E. Stelmach and Jean Requin, 647–656, Amsterdam: North Holland.
Bybee, Joan L. (1985) *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
Bybee, Joan L. (2000) The phonology of the lexicon: Evidence from lexical diffusion, in *Usage-based Models of Language*, ed. by Michael Barlow and Suzanne Kemmer, 65–85, Stanford, CA: CSLI Publications.
Bybee, Joan L. (2001) *Phonology and Language Use.* Cambridge: Cambridge University Press.
Bybee, Joan L. (2007) *Frequency of Use and the Organization of Language*. New York: Oxford University Press.
Bybee, Joan L., Paromita Chakraborti, Dagmar Jung, and Joanne Scheibman. (1998) Prosody and segmental effect: Some paths of evolution for word stress. *Studies in Language* 22, 267–314.
Bybee, Joan L. and Paul Hopper (eds.) (2001) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.
Bye, Patrik. (2006) Subtraction, optimization, and the combinatorial lexicon. Manuscript, University of Tromsø.
Bye, Patrik, Elin Sagulin, and Ida Toivonen. (2009) Phonetic duration, phonological quantity and prosodic structure in Inari Saami. *Phonetica* 66, 199–221.
Byrd, Dani. (1993) A phase window framework for articulatory timing. *Phonology* 13, 139–169.
Byrd, Dani and Elliot Saltzman. (2003) The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31, 149–180.
Béland, Renée and Carole Paradis. (1997) Principled syllabic dissolution in a primary progressive aphasia case. *Aphasiology* 11, 1171–1196.
Cahill, Mike. (2001) The unusual tone system of Awad Bing. Presented at the 75th Annual Meeting of the Linguistic Society of America, Washington DC.
Cahill, Mike. (2007) More universals of tone. SIL Electronic Working Papers.
Caisse, Michelle. (1982) Cross-linguistic differences in fundamental frequency pertubation induced by voiceless unaspirated stops. Master's thesis, University of California, Berkeley.
Calabrese, Andrea. (1985) Metaphony in Salentino. *Rivista di grammatica generativa* 9–10, 1–140.
Calabrese, Andrea. (1988) *Towards a Theory of Phonological Alphabets*. MIT dissertation.
Calabrese, Andrea. (1995) A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26, 373–463.
Calabrese, Andrea. (2005) *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton de Gruyter.
Calder, Jonathan. (1988) Unification and phonological analysis. Presented at the meeting of the Linguistics Association of Great Britain.
Calhoun, Sasha. (2006) *Information Structure and the Prosodic Structure of English: A Probabilistic Relationship*. University of Edinburgh dissertation.
Campbell, Lyle. (1980) The psychological and sociological reality of Finnish vowel harmony, in *Issues in Vowel Harmony*, ed. by Robert M. Vago, 245–270, Amsterdam: Benjamins.
Campbell, Lyle. (1986) Testing phonology in the field, in *Experimental Phonology*,

ed. by John J. Ohala and Jeri J. Jaeger, 163–186, Orlando, FL: Academic Press.

Campbell, Nick and Alan W. Black. (1997) Prosody and the selection of source units for concatenative synthesis, in *Progress in Speech Synthesis*, ed. by Jan P. H. van Santen, Richard W. Sproat, Joseph P. Olive and J. Hirschberg, 279–282, New York: Springer-Verlag.

Campos-Astorkiza, Rebeka. (2007) *Minimal Contrast and the Phonology-Phonetics Interaction*. University of Southern California dissertation.

Canalis, Stefano. (2007) A note on the influence of Louis Hjelmslev's suprasegmental phonology, in *16th International Congress of Phonetic Sciences*, 1369–1372.

Canalis, Stefano. (2009) Post-tonic vowel harmony in some dialects of central Italy: The role of prosodic structure, contrast, and consonants, in *Phonetics and Phonology: Interactions and Interrelations*, ed. by Marina Vigário, Sónia Frota, and M. João Freitas, Amsterdam: Benjamins.

Cantineau, Jean. (1949) *Principes de phonologie*. Paris: Klincksieck. Translation of Trubetzkoy (1939).

Cao, Jianfen and Ian Maddieson. (1992) An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics* 20, 77–92.

Caramazza, Alfonso. (1986) On drawing inferences about the structural of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition* 5, 41–66.

Cardoso, Walcir. (2001) Variation patterns in across-word regressive assimilation in Picard: An optimality theoretic account. *Language Variation and Change* 13, 305–341.

Carey, Susan and Elsa Bartlett. (1978) Acquiring a single new word. *Papers and Reports on Child Language Development* 15, 17–29.

Carlson, Rolf and Björn Granström. (1974) A phonetically oriented programming language for rule description of speech, in *Speech Communication Seminar, Stockholm, August 1–3, 1974*, volume 2, 245–253.

Carlson, Rolf and Björn Granström. (1999) Speech synthesis, in *Handbook of Phonetic Sciences*, ed. by William J. Hardcastle and John Laver, 768–788, Oxford: Blackwell.

Carson-Berndsen, Julie. (1998) *Time Map Phonology: Finite State Models and Event Logics in Speech Recognition*. Dordrecht: Kluwer.

Carstairs-McCarthy, Andrew. (1998) Phonological constraints on morphological rules, in *Handbook of Morphology*, ed. by Andrew Spencer and Arnold Zwicky, 144–148, Oxford: Blackwell.

Carter, A. and LouAnn Gerken. (2004) Do children's omissions leave traces? *Journal of Child Language* 31, 561–586.

Carter, Hazel. (1974) Negative structures in the syntactic tone-phrasing system of Kongo. *Bulletin of the School of Oriental and African Studies* 37, 29–40.

Carter, M. (2001) Arabic, in *Facts about The World's Languages: An Encyclopedia of the World's Major Languages, Past and Present*, ed. by Jane Garry and Carl Rubino, New York: The H. W. Wilson Company.

Carvalho, Joachim Brandão de. (2002) Formally-grounded phonology: From constraint-based theories to theory-based constraints. *Studia Linguistica* 56, 227–263.

Casali, Roderic F. (1995) Labial opacity and roundness harmony in Nawuri. *Natural Language and Linguistic Theory* 13, 649–663.

Casali, Roderic F. (1996) *Resolving Hiatus*. UCLA dissertation. Published by Garland Publishing, New York, 1998.

Casali, Roderic F. (2002) Nawuri ATR harmony in typological perspective.

Journal of West African Languages 29, 3–43.

Casali, Roderic F. (2003) [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7, 307–382.

Caspers, Johanneke. (2000) Experiments on the meaning of four types of single accent intonation patterns in Dutch. *Language and Speech* 43, 127–161.

Cassimjee, Farida and Charles W. Kisseberth. (1998) Optimal domains theory and Bantu tonology: A case study from IsiXhosa and Shingazidja, in *Theoretical Aspects of Bantu Tone*, ed. by Larry M. Hyman and Charles W. Kisseberth, 33–132, Stanford, CA: CSLI.

Cebrian, Juli. (2002) *Phonetic Similarity, Syllabification and Phonotactic Constraints in the Acquisition of a Second Language Contrast*. University of Toronto dissertation.

Cedergren, Henrietta, Hélène Perreault, François Poiré and Pascale Rousseau. (1990) L'accentuation québécoise: Une approche tonale. *Revue québécoise de linguistique* 19, 25–38.

Cedergren, Henrietta and David Sankoff. (1974) Variable rules: Performance as a statistical reflection of competence. *Language* 50, 333–355.

Cedergren, Henrietta and Louise Simoneau. (1985) La chute des voyelles hautes en français de Montréal: "As-tu entendu la belle syncope?", in *Les tendences dynamiques du français parlé à Montréal*, ed. by Monique Lemieux and Henrietta Cedergren, 57–144, Bibliothèque Nationale du Québec.

Chafe, Wallace. (1968) The ordering of phononological rules. *International Journal of American Linguistics* 34, 115–136.

Chafe, Wallace. (1970) *A Semantically Based Sketch of Onondaga*. Baltimore, MD: Waverly Press.

Chafe, Wallace. (1977) Accent and related phenomena in the Five Nations Iroqouis languages, in *Studies in Stress and Accent*, ed. by Larry M. Hyman, 169–181, Los Angeles: University of Southern California Linguistics Department.

Chafe, Wallace. (1996) Sketch of Seneca: An Iroquoian language. *Handbook of American Indian Linguistics* 17, 551–579.

Chambers, K., K. Onishi, and C. Fisher. (2003) Infants learn phonotactic regularities from brief auditory experience. *Cognition* 87, B69–B77.

Chan, Marjorie M. K. (1998) Sentence particles *je* and *jek* in Cantonese and their distribution across gender and sentence types, in *Engendering Communication: Proceedings of the 5th Berkeley Women and Language Conference*, ed. by Ashlee Bailey, Suzanne Wertheim, and Monica Corston-Oliver, 117–128.

Chan, Marjorie M. K. and Hongmao Ren. (1989) Wuxi tone sandhi: From last to first syllable dominance. *Acta Linguistica Hafniensia* 21, 35–64.

Chang, Nien-chuang T. (1958) Tone and intonation in the Chengtu dialect (Szechuan, China). *Phonetica* 2, 59–85.

Channon, Rachel. (2002) Beads on a string? Representations of repetition in spoken and sign languages, in *Modality and Structure in Signed and Spoken Languages*, ed. by Richard Meier, D. Quinto-Pozos, and K. Cormier, 65–87, Cambridge: Cambridge University Press.

Chao, Yuen Ren. (1930) A system of tone-letters. *Le Maître phonétique* 30, 24–27.

Chao, Yuen Ren. (1932) Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology, Academia Sinica* 4, 121–134.

Chao, Yuen Ren. (1934) On the non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4, 363–397.

Chao, Yuen Ren. (1968) *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.

Charette, Monik. (1990) Licence to govern. *Phonology* 7, 233–254.
Charette, Monik. (1991) *Conditions on Phonological Government*. Cambridge: Cambridge University Press.
Charette, Monik. (1992) Mongolian and Polish meet government licensing. *SOAS Working Papers in Linguistics and Phonetics* 2, 275–291.
Charette, Monik. (2003) Empty and pseudo-empty categories, in *Living on the Edge: 28 Papers in Honour of Jonathan Kaye*, ed. by Stefan Ploch, 465–480, Berlin: Mouton de Gruyter.
Charette, Monik. (2004) Defining the structure of Turkish words. *SOAS Working Papers in Linguistics* 13, 49–79.
Chase, Chris and Annette R. Jenner. (1993) Magnocellular visual deficits affect temporal processing of dyslexics. *Annals of the New York Academy of Sciences* 682, 326.
Chen, Matthew Y. (1970) Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.
Chen, Matthew Y. (1972) On the formal expression of natural rules in phonology. *Journal of Linguistics* 9, 209–383.
Chen, Matthew Y. (1987) The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4, 109–149.
Chen, Matthew Y. (2000) *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge: Cambridge University Press.
Chen, Yiya. (2008) Revisiting the phonetics and phonology of Shanghai tone sandhi, in *Proceedings of the Fourth Conference on Speech Prosody, 6–8 May 2008, Campinas, Brazil*, 253–256.
Chen, Yiya and Yi Xu. (2006) Production of weak elements in speech: Evidence from F0 patterns of neutral tone in Standard Chinese. *Phonetica* 63, 47–75.
Cherry, E. Colin, Morris Halle, and Roman Jakobson. (1953) Toward the logical description of languages in their phonemic aspect. *Language* 29, 34–43.
Chida, Shuntaro. (2001) On the tone of the Shanghai language. English translation by Yoshihiro Kawabata of Syuntarô Tida (2000), Domugo no tôn ni kansuru kôsatu. Master's thesis, Kyoto University.
Chin, S. and Daniel A. Dinnsen. (1992) Consonant clusters in disordered speech: Constraints and correspondence patterns. *Journal of Child Language* 19, 259–285.
Cho, Taehong and Peter Ladefoged. (1999) Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27, 207–229.
Cho, Young-mee Yu. (1990) *Parameters of Consonantal Assimilation*. Stanford University dissertation.
Chomsky, Noam. (1951) Morphophonemics of Modern Hebrew. Master's thesis, Philadelphia: University of Pennsylvania.
Chomsky, Noam. (1956) Three models for the description of language. *IEEE Transactions on Information Theory* 2, 113–124.
Chomsky, Noam. (1957) *Syntactic Structures*. The Hague: Mouton.
Chomsky, Noam. (1964) *Current Issues in Linguistic Theory*. The Hague: Mouton.
Chomsky, Noam. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam. (1970) Remarks on nominalization, in *Readings in English Transformational Grammar*, ed. by R. Jacobs and P. Rosenbaum, 184–221, Waltham, MA: Ginn and Co.
Chomsky, Noam. (1976) Conditions on rules of grammar. *Linguistic Analysis 2*, 303–351.
Chomsky, Noam. (1980) *Rules and Representations*. New York: Columbia University Press.
Chomsky, Noam. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. (2001) Derivation by phase, in *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1–52, Cambridge, MA: MIT Press.

Chomsky, Noam and Morris Halle. (1965) Some controversial questions in phonological theory. *Journal of Linguistics* 1, 97–138.
Chomsky, Noam and Morris Halle. (1968) *The Sound Pattern of English*. New York: Harper and Row.
Chomsky, Noam, Morris Halle, and Fred Lukoff. (1956) On accent and juncture in English, in *For Roman Jakobson: Essays on the Occasion of his 60th Birthday*, ed. by Morris Halle, 65–80, Mouton.
Chomsky, Noam and Howard Lasnik. (1977) Filters and control. *Linguistic Inquiry* 8, 425–504.
Chomsky, Noam and George A. Miller. (1958) Finite state languages. *Information and Control* 1, 91–112.
Church, Kenneth W. (1983) *Phrase Structure Parsing: A Method for Taking Advantage of Allophonic Constraints*. MIT dissertation.
Churchyard, Henry. (1989) Vowel reduction in Tiberian Biblical Hebrew as evidence for a sub-foot level of maximaly trimoraic metrical constituents, in *Coyote Papers 2, Proceedings of the Arizona Phonology Conference*, 1–19.
Churchyard, Henry. (1999) *Topics in Tiberian Biblical Hebrew Metrical Phonology and Prosodics*. University of Texas, Austin dissertation.
Churma, Don. (1985) *Arguments from External Evidence in Phonology*. New York: Garland Publishing.
Churma, Donald and Yili Shi. (1995) Glottal consonants and the "sonority" hierarchy, in *ESCOL 1995*.
Cinque, Gugliemo. (1993) A null theory of phrase and compound stress. *Linguistic Inquiry* 24, 239–237.
Clahsen, Harald, Fraibet Aveledo, and Iggy Roca. (2002a) The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language* 29, 591–622.
Clahsen, Harald, Peter Prufert, Sonja Eisenbeis, and Joana Cholin. (2002b) Strong stems in the German mental lexicon: Evidence from child language acquisition and adult processing, in *More than Words: Festschrift for Dieter Wunderlich*, ed. by I. Kaufmann and B. Stiebels, 91–112, Berlin: Akademie Verlag.
Clark, Mary Morris. (1978) *A Dynamic Treatment of Tone with Special Attention to the Tonal System of Igbo*. University of New Hampshire dissertation.
Clayards, Meghan. (2008) *The Ideal Listener: Making Optimal Use of Acoustic-Phonetic Cues for Word Recognition*. University of Rochester dissertation.
Clayards, Meghan, Michael K. Tanenhaus, Richard N. Aslin, and Robert A. Jacobs. (2008) Perception of speech reflects optimal use of probabilistic speech cues. *Cognition* 108, 804–809.
Clayton, Mary. (1976) The redundancy of morpheme structure conditions. *Language* 52, 296–313.
Clements, G. N. (1976) The autosegmental treatment of vowel harmony, in *Phonologica*, ed. by Wolfgang Dressler and O. Pfeiffer, 111–119, Innsbruck.
Clements, G. N. (1978) Tone and syntax in Ewe, in *Elements of Tone, Stress, and Intonation*, ed. by Donna Jo Napoli, 21–99, Washington, DC: Georgetown University Press.
Clements, G. N. (1979) The description of terraced-level tone languages. *Language* 55, 536–558.
Clements, G. N. (1980) *Vowel Harmony in Nonlinear Generative Phonology: An Autosegmental Model*. Bloomington: Indiana University Linguistics Club.
Clements, G. N. (1981) Akan vowel harmony: A nonlinear analysis. *Harvard Studies in Phonology* 2, 108–177.
Clements, G. N. (1985) The geometry of phonological features. *Phonology Yearbook* 2, 225–252.
Clements, G. N. (1990) The role of the sonority cycle in core syllabification, in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, ed. by John C. Kingston and Mary E. Beckman,

283–333, New York: Cambridge University Press.
Clements, G. N. (1991) Vowel height assimilation in Bantu languages, in *Proceedings of the Special Session on African Language Structures*, *BLS* 17, ed. by K. Hubbard, 25–64, Berkeley: University of California Press.
Clements, G. N. (2001) Representational economy in constraint based phonology, in *Distinctive Feature Theory, Phonology and Phonetics 2*, ed. by T. Alan Hall, 71–146, Berlin: Mouton de Gruyter.
Clements, G. N. (2003) Feature economy in sound systems. *Phonology* 20, 287–333.
Clements, G. N. and Kevin C. Ford. (1979) Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10, 179–210.
Clements, G. N. and John A. Goldsmith. (1984) *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris.
Clements, G. N. and Elizabeth V. Hume. (1995) The internal organization of speech sounds, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 245–306, Oxford: Blackwell.
Clements, G. N. and Samuel J. Keyser. (1983) *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.
Clements, G. N., Alexis Michaud, and Cédric Patin. (to appear) Do we need tone features?, in *Tones and Segments*, ed. by John A. Goldsmith, Elizabeth V. Hume and Leo Wetzels, Mouton.
Clements, G. N. and S. Osu. (2003) Ikwere nasal harmony in typological perspective, in *Typologie des langues d'Afrique et universaux de la grammaire*, ed. by Patrick Sauzet and Anne Zribi-Hertz, volume 2, 70–95, Paris: L'Harmattan.
Clements, G. N. and Engin Sezer. (1982) Vowel and consonant disharmony in Turkish, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 2, 213–255, Dordrecht: Foris.
Clocksin, Wiliam F. and Christopher S. Mellish. (1981) *Programming in Prolog*. Berlin: Springer-Verlag.
Cockerill, MaryAnn. (1998) The formation of hypocoristics in Arabic. Manuscript, Harvard University.
Coetzee, Andries W. (2004) *What it Means to be a Loser: Non-Optimal Candidates in Optimality Theory*. University of Massachusetts, Amherst dissertation.
Coetzee, Andries W. (2006) Variation as accessing "non-optimal" candidate. *Phonology* 23, 337–385.
Coetzee, Andries W. (2008) Grammaticality and ungrammaticality in phonology. *Language* 84, 218–257.
Coetzee, Andries W. (2009a) An integrated grammatical/non-grammatical model of phonological variation, in *Current Issues in Linguistic Interfaces*, vol. 2, ed. by Young-Se Kang, Jong-Yurl Yoon, Hyunkyung Yoo, Sze-Wing Tang, Yong-Soon Kang, Youngjun Jang, Chul Kim, Hyoung-Ae Kim, and Hye Kyung Kang, 267–294, Seoul: Hankookmunwhasa.
Coetzee, Andries W. (2009b) Phonological variation and lexical frequency, in *NELS 38: Proceedings of the 38th Meeting of the North East Linguistics Society*, vol. 1, ed. by Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, Amherst, MA: GLSA. Available on Rutgers Optimality Archive, ROA-952.
Coetzee, Andries W., Susan Lin and Rigardt Pretorius. (2007) Post-nasal devoicing in Tswana, in *Proceedings of the International Congress of Phonetic Sciences*, ed. by Jürgen Trouvain and William J. Barry, volume 16, 861–864, Saarbrücken: Pirrot.
Coetzee, Andries W. and Joe Pater. (2008) Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26, 289–337.
Cohen, Antonie and Johan 't Hart. (1967) On the anatomy of intonation. *Lingua* 19, 177–192.

Cohen, P. and R. Mercer. (1974) The phonological component of an automatic speech-recognition system, in *Speech Recognition: Invited Papers Presented at the 1974 IEEE Symposium*, ed. by D. R. Reddy, 275–320, New York: Academic Press.

Cohn, Abigail. (1989) Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7, 167–216.

Cohn, Abigail. (1990) *Phonetic and Phonological Rules of Nasalization*. UCLA dissertation.

Cohn, Abigail. (1992) The consequences of dissimilation in Sudanese. *Phonology* 9, 199–220.

Cohn, Abigail. (1993a) Nasalisation in English: Phonology or phonetics. *Phonology* 10, 43–81.

Cohn, Abigail. (1993b) The status of nasalized continuants, in *Nasals, Nasalization, and the Velum. Phonetics and Phonology, Vol. 5*, ed. by Marie K. Huffman and Rena Krakow, 329–367, San Diego, CA: Academic Press.

Cohn, Abigail. (1993c) A survey of the phonology of the feature [±nasal]. *Working Papers of the Cornell Phonetics Laboratory* 8, 141–203.

Cole, Jennifer. (1987) *Planar Phonology and Morphology*. MIT dissertation.

Cole, Jennifer and John Coleman. (1993) No need for cyclicity in generative phonology, in *Papers from the 28th Regional Meeting of the Chicago Linguistic Society: Parasession on the Cycle in Linguistic Theory*, ed. by J. M. Denton, G. P. Chan, and C. P. Canakis, 36–50, Chicago: Chicago Linguistic Society.

Cole, Jennifer, Gary Dell and Alina Khasanova. (2002) Evidence of a production basis for front/back vowel harmony. Poster presented at the 8th Conference on Laboratory Phonology.

Cole, Jennifer and José Hualde. (1998) The object of lexical acquisition: A UR-free model, in *Papers from the 34th Annual Meeting of the Chicago Linguistic Society: The Panels*, 447–458, Chicago: Chicago Linguistic Society.

Cole, Jennifer and Charles Kisseberth. (1995a) Nasal harmony in optimal domains theory. Cognitive Science Technical Report UIUC-BI-CS-95-02, Beckman Institute, University of Illinois.

Cole, Jennifer and Charles Kisseberth. (1995b) An optimal domains theory of vowel harmony, in *Formal Linguistics Society of Mid-America V*, 101–114, Urbana, University of Illinois.

Cole, Jennifer and Charles Kisseberth. (1995c) Restricting multi-level constraint evaluation: Opaque rule interaction in Yawelmani vowel harmony, in *Proceedings of the South Western Optimality Theory Workshop*, ed. by Keiichiro Suzuki and Dirk Elzinga, 18–38, University of Arizona Coyote Papers.

Cole, Jennifer and Loren Trigo. (1988) Parasitic harmony, in *Features, Segmental Structure and Harmony Processes*, ed. by Harry van der Hulst and Norval Smith, 19–38, Dordrecht: Foris.

Cole, Ronald A., Max Coltheart, and Fran Allard. (1974) Memory of a speaker's voice: Reaction time to same- or different-voiced letters. *The Quarterly Journal of Experimental Psychology* 26, 1–7.

Coleman, John. (1992) "Synthesis-by-rule" without segments or rewrite-rules, in *Talking Machines: Theories, Models and Designs*, ed. by G. Bailly, C. Benoît, and T. R. Sawallis, 43–60, Elsevier.

Coleman, John. (1996) Declarative syllabification in Tashlhiyt Berber, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks, volume 1, 177–218, Salford: European Studies Research Institute, University of Salford.

Coleman, John. (1998) *Phonological Representations: Their Names, Forms and Powers*. Cambridge: Cambridge University Press.

Coleman, John. (2000) Candidate selection. *The Linguistic Review* 17, 167–179.
Coleman, John. (2002) Phonetic representations in the mental lexicon, in *Phonetics, Phonology, and Cognition*, ed. by Jacques Durand and Bernard Laks, 96–130, Oxford: Oxford University Press.
Coleman, John. (2003) Discovering the acoustic correlates of phonological contrasts. *Journal of Phonetics* 31, 351–372.
Coleman, John. (2005) *Introducing Speech and Language Processing*. Cambridge: Cambridge University Press.
Coleman, John. (2006) Declarative approaches to phonology, in *Encyclopedia of Language and Linguistics*, ed. by Keith Brown, volume 3, 374–377, Oxford: Elsevier.
Coleman, John and Janet B. Pierrehumbert. (1997) Stochastic phonological grammars and acceptability, in *Computational Phonology: Third meeting of the ACL Special Interest Group in Computational Phonology*, 49–56, Somerset, NJ: Association for Computational Linguistics.
Collier, René Collier. (1975) Physiological correlates of intonation patterns. *Journal of the Acoustical Society of America*, 58, 249–255.
Conlin, Kimberly E., Gene Mirus, Claude Mauk and Richard P. Meier. (2000) Acquisition of first signs: Place, handshape, and movement, in *The Acquisition of Linguistic Representation by Eye*, ed. by C. Chamberlain, J. P. Morford, and R. Mayberry, 51–69, Mahwah, NJ: Lawrence Erlbaum Associates.
Connelly, M. (1984) *Basotho Children's Acquisition of Noun Morphology.* University of Essex dissertation.
Connine, Cynthia M. and Eleni Pinnow. (2006) Phonological variation in spoken word recognition: Episodes and abstractions. *The Linguistic Review* 23, 235–245.
Cook, E.-D. (1976) A Phonological Study of Chilcotin and Carrier. Technical Report E10-75-8, National Museum of Man, Ottawa.
Cook, E.-D. (1983) Chilcotin flattening. *Canadian Journal of Linguistics* 28, 123–132.
Cook, E.-D. (1993) Chilcotin flattening and autosegmental phonology. *Lingua* 91, 149–174.
Cooper, William E. and Jeanne Paccia-Cooper. (1980) *Syntax and Speech*. Cambridge, MA: Harvard University Press.
Corbett, Greville G., Norman M. Fraser, and Scott McGlashan (eds.) (1993) *Heads in Grammatical Theory*. Cambridge University Press.
Corina, David. (1990) Assessing the role of sonority in syllable structure: Evidence from a visual-gestural language. *Proceedings from the 26th Annual Meeting of the Chicago Linguistics Society*, Vol. 2: Parasession on the Syllable in Phonetics and Phonology, 33–44. Chicago: Chicago Linguistics Society.
Côté, Marie-Hélène and Geoffrey Stewart Morrison. (2007) The nature of the schwa-zero alternation in French clitics: Experimental and non-experimental evidence. *Journal of French Language Studies* 17: 159–186.
Coulter, Geoffrey. (1982) On the nature of ASL as a monosyllabic language. Presented at the annual meeting of the Linguistic Society of America, San Diego, CA.
Coussé, Evie, Steven Gillis, Hanne Kloots, and Marc Swerts. (2004) The influence of the labeller's regional background on phonetic transcriptions: Implications for the evaluation of spoken language resources. *Proceedings of the 4th International Conference on Language Resources and Evaluation*, 1447–1450, Paris: ELDA.
Cowan, Nelson and Lewis Leavitt. (1982) Talking backward: Exceptional speech play in late childhood. *Journal of Child Language* 9, 481–495.

Cowan, Nelson, Lewis A. Leavitt, Dominic Massaro, and Raymond Kent. (1982) A fluent backward talker. *Journal of Speech and Hearing Research* 25, 48–53.

Craik, Fergus I. M. and Kim Kirsner. (1974) The effect of speaker's voice on word recognition. *The Quarterly Journal of Experimental Psychology* 26, 274–284.

Crasborn, Onno. (2001) *Phonetic Implementation of Phonological Categories in Sign Language of the Netherlands*. Utrecht: LOT (Netherlands Graduate School of Linguistics).

Creel, Sarah C., Elissa L. Newport, and Richard N. Aslin. (2004) Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 1119–1130.

Creissels, D. and C. Grégoire. (1993) La notion de ton marqué dans l'analyse d'une opposition tonale binaire: Le cas du mandingue. *Journal of African Languages and Linguistics* 14, 107–154.

Cristófaro-Silva, Thais. (1992) *Nuclear Phenomena in Brazilian Portuguese*. School of Oriental and African Studies, University of London dissertation.

Crosswhite, Katherine. (1999) Intra-paradigmatic homophony avoidance in two dialects of Slavic, in *UCLA Working Papers in Linguistics: Papers in Phonology*, ed. by M. K. Gordon, 48–67, Los Angeles: UCLA Department of Linguistics.

Crosswhite, Katherine. (2001) *Vowel Reduction in Optimality Theory*. London: Routledge.

Crosswhite, Katherine. (2004) Vowel reduction, in *Phonetically Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 191–231, Cambridge: Cambridge University Press.

Crowhurst, Megan. (1991) *Minimality and Foot Structure in Metrical Phonology and Prosodic Morphology*. University of Arizona dissertation.

Crowhurst, Megan and Mark Hewitt. (1994) Directional footing, degeneracy, and alignment. ROA-65.

Crowhurst, Megan and Lev Michael. (2005) Iterative footing and prominence-driven stress in Nanti. *Language* 81, 47–95.

Cruttenden, A. (1970) A phonetic study of babbling. *International Journal of Language and Communication Disorders* 5, 110–117.

Cucchiarini, Catia (1993) *Phonetic Transcription: A Methodological and Empirical Study*. Nijmegen: Catholic University Nijmegen.

Culbertson, Jennifer and Steven Gross. (2009) Are linguists better subjects? *British Journal for the Philosophy of Science* 60, 721–736.

Cummins, Robert. (1983) *The Nature of Psychological Explanation*. Cambridge MA: MIT Press.

Curtin, Suzanne and Kie Zuraw. (2002) Explaining constraint demotion in a developing system, in *BUCLD 26: Proceedings of the 26th Annual Boston University Conference on Language Development*, ed. by Anna H.-J. Do, Laura Dominguez, and Aimee Johansen, Somerville, MA: Cascadilla Press.

Curtis, Emily. (2003) *Geminate Weight: Case Studies and Formal Models*. University of Washington dissertation.

Cutler, Anne, Jacques Mehler, Dennis Norris, and Jean Segui. (1986) The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25, 385–400.

Cuxac, Christian and Marie-Anne Sallandre. (2007) Iconicity and arbitrariness in French Sign Language–Highly iconic structures, degenerated iconicity and diagrammatic iconicity, in *Verbal and Signed Languages, Comparing Structures, Constructs, and Methodologies*, ed. by E. Pizzuoto, P. Pietrandrea, and R. Simone, 13–34, Berlin: Mouton de Gruyter.

Cyran, Eugeniusz. (1996) Licensing properties of nuclei and principle ranking in Irish. *The Linguistic Review* 13, 1–31.

Cyran, Eugeniusz (ed.) (1998) *Structure and Interpretation: Studies in Phonology*. Folium.

Cyran, Eugeniusz. (2006) Review of "*A Lateral Theory of Phonology*," by Tobias Scheer. *The Linguistic Review* 23, 505–542.

Cyran, Eugeniusz. (2010) *Complexity Scales and Licensing Strength in Phonology*. New York/Berlin: Mouton de Gruyter.

Cyran, Eugeniusz and Edmund Gussman. (1999) Consonant clusters and governing relations: Polish initial consonant sequences, in *The Syllable: Views and Facts*, ed. by Harry van der Hulst and Nancy A. Ritter, Berlin: Mouton de Gruyter.

Czaykowska-Higgins, Ewa. (1993) Cyclicity and stress in Moses-Columbia Salish (Nxa'amxcin). *Natural Language and Linguistic Theory* 11, 197–278.

Daelemans, Walter and Antal van den Bosch. (1992) Generalization performance of back-propagation learning on a syllabification task, in *Proceedings of the 3rd Twente Workshop on Language Technology*, 27–38.

Daelemans, Walter, Antal Van den Bosch, and J. Zavrel. (1999) Forgetting exceptions is harmful in language learning. *Machine Learning* 34, 11–41.

Daelemans, W. and A. Van den Bosch. (2005) *Memory-based, Language Processing*. Cambridge: Cambridge University Press.

Daelemans, Walter, Jakub Zavrel, Ko Van der Sloot and Antal Van den Bosch. (2000) TiMBL: Tilburg Memory Based Learner Reference Guide 3.0. Technical Report 00-01, Computational Linguistics Tilburg University.

Dainora, A. (2001) Eliminating downstep in prosodic labeling of American English, in *Proceedings of the ISCA Tutorial and Research Workshop (ITRW) on Prosody in Speech Recognition and Understanding*, 41–46, Red Bank.

Daly, John P. and Larry M. Hyman. (2007) On the representation of tone in Peñoles Mixtec. *International Journal of American Linguistics* 73, 165–208.

Danesi, Marcel. (1985) *Loanwords and Phonological Methodology*. Ville LaSalle, Quebec: Didier.

Dankovicová, Jana, Paula West, John Coleman, and Andrew Slater. (1998) Phonotactic grammaticality is gradient. Poster at the 6th Laboratory Phonology Conference, University of York, York, UK. Retrieved June 1, 2007, from http://www.phon.ox.ac.uk/jcoleman/LabPhon-poster-paper.pdf.

Darden, Bill. (1978) Local ordering vs. multiple application of phonological rules, in *Papers from the 14th Annual Meeting of the Chicago Linguistic Society*, 65–75, Chicago: Chicago Linguistic Society.

Darwin, C. (1877) A biographical sketch of an infant. *Mind*, 285–294.

Dauer, Rebecca M. (1980) The reduction of unstressed high vowels in Modern Greek. *Journal of the International Phonetic Association* 10, 17–27.

Davidson, Lisa. (2006a) Phonology, phonetics, or frequency: Influence on the production of non-native sequences. *Journal of Phonetics* 34, 104–137.

Davidson, Lisa. (2006b) Phonotactics and articulatory coordination interact in phonology: Evidence from non-native production. *Cognitive Science* 30, 837–862.

Davidson, Lisa, Peter Jusczyk, and Paul Smolensky. (2004) The initial and final states: Theoretical and experimental explorations of richness of the base, in *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 321–368, Cambridge: Cambridge University Press.

Davidson, Lisa. (2007) The relationship between the perception of non-native

phonotactics and loanword adaptation. *Phonology* 24, 261–286.
Davies, H. J. (1980) *Kobon Phonology*. Canberra: Australian National University.
Davis, Donald R. (1969) The distinctive features of the Wantoat phonemes. *Linguistics* 47, 5–17.
Davis, Stuart. (1985) *Topics in Syllable Geometry*. University of Arizona dissertation.
Davis, Stuart. (1988) Syllable onsets as a factor in stress rules. *Phonology* 5, 1–19.
Davis, Stuart. (1990) The onset as a constituent of the syllable, in *Papers from, the 34th Annual Meeting of the Chicago Linguistic Society*, ed. by Michael Ziolkowski, Manuela Noske and Karen Deaton, volume 2: *The Parasession on the Syllable in Phonetics and Phonology*, 71–79, Chicago: Chicago Linguistic Society.
Davis, Stuart. (1994) Geminate consonants in moraic phonology, in *Proceedings of the 13th Annual Meeting of the West Coast Conference on Formal Linguistics*, 32–45.
Davis, Stuart. (1995a) Emphasis spread in Arabic and Grounded Phonology. *Linguistic Inquiry* 26, 465–498.
Davis, Stuart. (1995b) Weight properties of partial geminates. Presented at a meeting of the Linguistic Society of America, New Orleans.
Davis, Stuart. (1999a) On the moraic representation of underlying geminates: Evidence from prosodic morphology, in *The Prosody-Morphology Interface*, ed. by Harry van der Hulst, René Kager, and Wim Zonneveld, 39–61, Cambridge: Cambridge University Press.
Davis, Stuart. (1999b) On the representation of initial geminates. *Phonology* 16, 93–104.
Davis, Stuart. (2003) The controversy over geminates and syllable weight, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, 77–98, Cambridge: Cambridge University Press.
Davis, Stuart. (2004) Phonetics vs. phonology: The word final /s/ problem in Korean loanword adaptation. Presented at the 12th Manchester Phonology Meeting, Manchester.
Davis, Stuart. (2006) Syllable constituents, in *The Encyclopedia of Language and Linguistics*, volume 12, 326–328.
Davis, Stuart and Karen Baertsch. (2005) The diachronic link between onset clusters and codas, in *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*, 397–408.
Davis, Stuart and Michael Hammond. (1995) On the status of onglides in American English. *Phonology* 12, 159–182.
Davis, Stuart and Gina Torretta. (1998) An optimality-theoretic account of compensatory lengthening and geminate throwback in Trukese, in *Proceedings of the 28th Annual Meeting of the North East Linguistic Society*, 111–125.
Davis, Stuart and Isao Ueda. (2006) Prosodic vs. morphological mora augmentation. *Lexicon Forum* 2, 121–143.
Davis, Stuart and Bushra Zawaydeh. (2001) Arabic hypocoristics and the status of the consonantal root. *Linguistic Inquiry* 32, 512–520.
Day, Ruth S. (1968) *Fusion in Dichotic Listening*. Stanford University dissertation.
Day, Ruth S. (1976) Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? Technical report, New Haven, CT: Haskins Laboratories.
de Boer, Bart. (2001) *The Origins of Vowel Systems*. Oxford: Oxford University Press.
de Boer, Bart and Patricia Kuhl. (2003) Investigating the role of infant-directed speech with a computer model. *Acoustic Research Letters On-line* 4, 129–134.
de Jong, Kenneth. (1998) Stress-related variation in the articulation of coda alveolar stop: Flapping revisited. *Journal of Phonetics* 26, 283–310.

de Jong, Kenneth. (2000) Attention modulation and the formal properties of stress systems, in *Papers from the 36th Annual Meeting of the Chicago Linguistic Society*, 71–91, Chicago: Chicago Linguistic Society.

de Jong, Kenneth, Mary E. Beckman, and Jan Edwards. (1993) The interplay between prosodic structure and coarticulation. *Language and Speech* 36, 197–212.

de Lacy, Paul. (2002a) *The Formal Expression of Markedness*. University of Massachusetts dissertation.

de Lacy, Paul. (2002b) The interaction of tone and stress in Optimality Theory. *Phonology* 19, 1–32.

de Lacy, Paul. (2003) Constraint universality and prosodic phrasing in Maori, in *Papers in Optimality II*, ed. by Angela Carpenter, Andries Coetzee, and Paul de Lacy, 59–79, Amherst, MA: GLSA Publications.

de Lacy, Paul. (2004) Markedness conflation in Optimality Theory. *Phonology* 21, 145–199.

de Lacy, Paul. (2006) Transmissibility and the role of the phonological component. *Theoretical Linguistics* 32, 185–196.

de Lacy, Paul and John Kingston. (2006) Synchronic explanation. Manuscript, Rutgers University and University of Massachusetts, Amherst.

de Marcken, Carl. (1996) *Unsupervised Language Acquisition*. MIT dissertation.

de Pijper, Jan Roelof. (1983) *Modelling British English Intonation*. Dordrecht: Foris.

de Rijk, Rudolf. (1970) Vowel interaction in Bizcayan Basque. *Fontes Linguae Vasconum* 2, 149–167.

de Vaan, Laura, Rob Schreuder, and R. Harald Baayen. (2007) Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2, 1–23.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux, and Ariel Gout. (2000) Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12, 635–647.

Dehé, Nicole. (2009) Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45, 569–615.

Delais-Roussarie, Elisabeth. (1995) *Pour une approche parallèle de la structure prosodique: Étude de l'organisation prosodique et rythmique de la phrase française*. Université de Toulouse – Le Mirail: dissertation.

Delattre, Pierre. (1965) *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report*. Philadelphia: Chilton Books and Groos Verlag.

Delattre, Pierre. (1981) An acoustic and articulatory study of vowel reduction in four languages, in *Studies in Comparative Phonetics*, ed. by Pierre Delattre, 63–93, Heidelberg: Julius Groos Verlag.

Dell, François. (1973) *Les Règles et les sons*. Orleabs: Imprimerie Nouvelle. Trans. 1980 by Catherine Cullen as *Generative Phonology and French Phonology*. Cambridge: Cambridge University Press.

Dell, François. (1981) On the learnability of optional phonological rules. *Linguistic Inquiry* 12, 31–37.

Dell, François. (1995) Consonant clusters and phonological syllables in French. *Lingua*, 95, 5–26.

Dell, François and Mohamed Elmedlaoui. (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7, 105–130.

Dell, François and Mohamed Elmedlaoui. (1988) Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics* 10, 1–17.

Dell, François and Mohamed Elmedlaoui. (1998) Nonsyllabic transitional vocoid in Imdlawn Tashlhiyt, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks,

volume 1, 217–244, Salford: European Studies Research Institute, University of Salford.

Dell, François and Mohamed Elmedlaoui. (2002) *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.

Dell, Gary. (1986) A spreading-activation theory of retrieval in sentence production. *Psychologial Review* 93, 3, 283–381.

Dell, Gary, Lisa Burger, and William R. Svec. (1997) Language production and serial order: A functional analysis and a model. *Psychological Review* 104, 123–47.

Dell, Gary, Kristopher D. Reed, David R. Adams, and Antje S. Meyer. (2000) Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26, 1355–1367.

Demuth, Katherine. (1993) Issues in the acquisition of the Sesotho tonal system. *Journal of Child Language* 20, 275–301.

Demuth, Katherine. (1994) On the "underspecification" of functional categories in early grammars, in *Syntactic Theory and First Language Acquisition: Cross-linguistic Perspectives*, ed. by B. Lust, M. Suñer, and J. Whitman, volume 1, 119–134, Hillsdale, NJ: Lawrence Erlbaum Associates.

Demuth, Katherine. (1995) Markedness and the development of prosodic structure, in *Proceedings of the 25th Annual Meeting of the North East Linguistic Society*, 13–26, Amherst, MA: GLSA, University of Massachusetts.

Demuth, Katherine. (1996) Alignment, stress and parsing in early phonological words, in *Proceedings of the UBC International Conference on Phonological Acquisition*, 113–124, Somerville, MA: Cascadilla Press.

Demuth, Katherine. (1997) Multiple optimal outputs in acquisition. *University of Maryland Working Papers in Linguistics* 5, 53–71.

Demuth, Katherine. (2001) Prosodic constraints on morphological development, in *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, ed. by J. Weissenborn and B. Höhle, 3–21, Amsterdam: Benjamins.

Demuth, Katherine. (2006) Crosslinguistic perspectives on the development of prosodic words. *Language and Speech* 49, 129–297.

Demuth, Katherine, J. Culbertson, and J. Alter. (2006) Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech* 49, 137–174.

Demuth, Katherine and D. Ellis. (2009) Revisiting the acquisition of Sesotho noun class prefixes, in *Crosslinguistic Approaches to the Study of Language: Festschrift for Dan Slobin*, ed. by J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, and S. Ozçalikan, 93–104, Abingdon, UK: Routledge.

Demuth, Katherine and E. Fee. (1995) Minimal prosodic words in early phonological development. Unpublished manuscript.

Demuth, Katherine and Mark Johnson. (2003) Truncation to subminimal words in early French. *Canadian Journal of Linguistics* 48, 211–242.

Demuth, Katherine and M. Kehoe. (2006) The acquisition of word-final clusters in French. *Journal of Catalan Linguistics* 5, 59–81.

Demuth, Katherine and E. McCullough. (2009a) The acquisition of clusters in French. *Journal of Child Language* 36, 425–448.

Demuth, Katherine and E. McCullough. (2009b) The prosodic (re)organization of children's early English articles. *Journal of Child Language* 36, 173–200.

Demuth, Katherine, M. Patrolia, J. Y. Song, and M. Masapollo. (forthcoming) The development of articles in children's early Spanish: Prosodic interactions

between lexical and grammatical form, in *First Language*.
Demuth, Katherine and A. Tremblay. (2008) Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language* 35, 99–127.
Denes, P. and J. Milton-Williams. (1962) Further studies in intonation. *Language and Speech* 5, 1–14.
Denwood, Ann. (2006) A template for Turkish, full circle: templates or syllables. *The Linguistic Review* 23, 465–504.
Déreau, Léon. (1955) *Cours de kikongo*. Namur: A. Wesmael-Charlier.
Derwing, Bruce and Roberto de Almeida. (2009) Non-chronometric experiments in linguistics, in *Experimental and Quantitative Linguistics*, ed. by D. Eddington, Munich: Lincom.
Derwing, Bruce, Maureen Dow, and Terrance Nearey. (1988) Experimenting with syllable structure, in *Proceedings of the Eastern States Conference on Linguistics*, volume 5, 83–94.
Derwing, Bruce L. and William J. Baker. (1986) Assessing morphological development, in *Language Acquisition: Studies in First Language Development*, ed. by P. Fletcher and M. Garman, 326–338, Cambridge: Cambridge University Press, 2nd edition.
Dettweiler, Stephen. (2000) Vowel harmony and neutral vowels in C'Lela. *The Journal of West African Languages* 28, 3–18.
Deville, G. (1891) Notes sur le développement du langage II. *Revue de linguistique et de philologie comparée* 24, 10–42, 128–143, 242–257, 300–320.
Dewaele, Jean-Marc. (2001) Activation or inhibition: The interaction of L1, L2 and L3 on the language mode continuum, in *Cross-Linguistic Influence in Third Language Acquisition*, ed. by Jasone Cenoz, Britta Hufeisen, and Ulrike Jessner, 69–89, Clevedon: Multilingual Matters.
DiCanio, Christian. (2006) On non-optimal laryngeal timing: The case of Trique, in *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society*.
Diehl, Randy and Katherine Kolodzey. (1981) Spaka: A private language. *Language* 57, 406–424.
Dienes, Péter. (2000) VC Phonology: A theory of consonant lenition and phonotactics. Master's thesis, Eötvös Loránd University, Budapest.
Diffloth, Gerard. (1976) Expressives in Semai, in *Austroasiatic Studies*, ed. by P. Jenner *et al.*, Honolulu, HI: The University Press of Hawaii.
Dijkstra, Ton. (2003) Lexical processing in bilinguals and multilinguals, in *The Multilingual Lexicon*, ed. by Jasone Cenoz, Britta Hufeisen, and Ulrike Jessner, 11–26, Dordrecht: Kluwer Academic Publishers.
Dilley, Laura C. and Mark A. Pitt. (2007) A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *The Journal of the Acoustical Society of America*, 122, 2340–2353.
Dimmendaal, Gerrit. (1983) *The Turkana Language*. Dordrecht: Foris.
D'Imperio, Mariapaola, Gorka Elordieta, Sónia Frota, Pilar Prietó, and Marina Vigário. (2005) Intonational phrasing in Romance: The role of prosodic and syntactic structure, in *Prosodies*, ed. by Sónia Frota, Marina Vigário, and M. João Freitas, 59–97, Berlin: Mouton de Gruyter.
D'Imperio, Mariapaola, and Barbara Gili Fivela. (2003) How many levels of phrasing? Evidence from two varieties of Italian, in *Papers in Laboratory Phonology V*, ed. by John Local, Richard Ogden, and Rosalind Temple, 130–144, Cambridge: Cambridge University Press.
Dinnsen, Daniel. (1972) *General Constraints on Phonological Rules*. University of Texas Austin dissertation.
Dinnsen, Daniel. (1979) Atomic phonology, in *Current Approaches to Phonological*

Theory, ed. by Daniel A. Dinnsen, 31–49, Bloomington, IN: Indiana University Press.
Dinnsen, Daniel (1985) A re-examination of phonological neutralization. *Journal of Linguistics* 21, 265–279.
Dinnsen, Daniel and Jessica A. Barlow. (1998) On the characterization of a chain shift in normal and delayed phonological acquisition. *Journal of Child Language* 25, 61–94.
Dinnsen, Daniel A. and Jan Charles-Luce. (1984) Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12, 49–60.
Dinnsen, Daniel A. and Ashley W. Farris-Trimble. (2008) An opacity-tolerant conspiracy in phonological acquisition, in *Phonological Opacity Effects in Optimality Theory, Indiana University Working Papers in Linguistics 6*, ed. by Ashley W. Farris-Trimble and Daniel A. Dinnsen, 99–118, Bloomington IN: IULC.
Dinnsen, Daniel and Judith Gierut. (2007) *Optimality Theory: Phonological Acquisition and Disorders*. London: Equinox Publishing.
Di Paolo, Marianna. (1988) Pronunciation and categorization in sound change, in *Linguistic Change and Contact: Proceedings of the 16th Annual Conference on New Ways of Analysing Variation in Language*, ed. by K. Ferrara, B. Brown, K. Walters, and J. Baugh, 84–92, University of Texas.
Disner, Sandra. (1983) Vowel quality: The relation between universal and language-specific factors, in *UCLA Working Papers in Phonetics*, volume 58, UCLA.
Dixon, Robert Malcolm Ward. (1977) *A Grammar of Yidiny*. Cambridge: Cambridge University Press.
Dobashi, Yoshihito. (2003) *Phonological Phrasing and Syntactic Derivation*. Cornell University dissertation.
Dobrovolsky, Michael. (1999) The phonetics of Chuvash stress: implications for phonology, in *Proceedings of the 14th International Conference of Phonetic Sciences*, 539–542.
Docherty, Gerard J. (1989) *An Experimental Phonetic Study of the Timing of Voicing in English Obstruents*. University of Edinburgh dissertation.
Dohlus, Katrin. (2005) Phonetics or phonology: Asymmetries in loanword adaptations – French and German mid-front rounded vowels in Japanese. *ZAZ Papers in Linguistics* 42, 117–135.
Doke, Clement. (1938) *Textbook of Lamba Grammar*. Johannesburg: Witwatersrand University Press.
Dolgopolsky, Aaron. (1977) Emphatic consonants in Semitic. *Israel Oriental Studies* 7, 1–13.
Donegan, Patricia Jane and David Stampe. (1979) The study of natural phonology, in *Current Approaches to Phonological Theory*, ed. by Daniel A. Dinnsen, 126–173, Bloomington, IN: Indiana University Press.
Donohue, Mark. (1997) Tone in New Guinea languages. *Linguistic Typology* 1, 347–386.
Donovan, R. E. (1996) *Trainable Speech Synthesis*. University of Cambridge dissertation.
dos Santos, C. (2007) *Développement phonologique en français langue maternelle. Une étude de cas*. Université Lumière Lyon dissertation.
Downing, Bruce. (1970) *Syntactic Structure and Phonological Phrasing in English*. University of Texas, Austin dissertation.
Downing, Bruce. (1973) Parenthetical rules and obligatory phrasing. *Papers in Linguistics* 6, 108–128.
Downing, Laura J. (1990) Local and metrical shift in Nguni. *Studies in African Linguistics* 21, 261–317.
Downing, Laura J. (1998) On the prosodic misalignment of onsetless syllables. *Natural Language and Linguistic Theory* 16, 1–52.
Downing, Laura J. (2004) The emergence of the marked: Tone in some African

reduplicative systems, in *Studies on Reduplication*, ed. by B. Hurch, 89–110, Berlin: Mouton de Gruyter.
Downing, Laura J. (2005a) Jita causative doubling provides optimal paradigms, in *Paradigms in Phonological Theory*, ed. by Laura Downing, T. Alan Hall, and R. Raffelsiefen, 122–144, Oxford: Oxford University Press.
Downing, Laura J. (2005b) On the ambiguous segmental status of nasals in homorganic NC sequences, in *The Internal Organization of Phonological Segments*, ed. by Marc van Oostendorp and Jeroen van de Weijer, 183–215, Berlin: Mouton de Gruyter.
Downing, Laura J. (2006) *Canonical Forms in Prosodic Morphology*. Oxford: Oxford University Press.
Downing, Laura J. (2007) African languages and the typology of focus prosody, in *Workshop on the Typology of African Languages*, Paris.
Downing, Laura J. (2011) The prosody of "dislocation" in selected Bantu languages. *Lingua* 121, 772–768.
Dresher, B. Elan. (1980) Metrical structure and secondary stress in Tiberian Hebrew, in *Brown Working Papers in Linguistics 4*, ed. by C. Chapin, 24–37, Rhode Island, RI: Brown University.
Dresher, B. Elan. (1981) On the learnability of abstract phonology, in *The Logical Problem of Language Acquisition*, ed. by C. L. Baker and John J. McCarthy, 188–210, Cambridge, MA: MIT Press.
Dresher, B. Elan. (2004) On the acquisition of phonological representations, in *Proceedings of 1st Workshop on Psycho-Computational Models of Human Language Acquisition*, 41–48.
Dresher, B. Elan. (2009) *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
Dresher, B. Elan and Harry van der Hulst. (1998) Head-dependent asymmetries in prosodic phonology: Visibility and complexity. *Phonology* 15, 317–52.
Dresher, B. Elan and Jonathan Kaye. (1990) A computational learning model for metrical phonology. *Cognition* 20, 421–451.
Dresher, B. Elan, Glyne Piggott, and Keren Rice. (1994) Contrast in phonology: Overview, in *Toronto Working Papers in Linguistics 13*, iii–xvii, University of Toronto.
Dressler, Wolfgang. (1985) *Morphonology, the Dynamics of Derivation*. Ann Arbor, MI: Karoma Publishers.
Duanmu, San. (1990) *A Formal Study of Syllable, Tone, Stress, and Domain in Chinese Languages*. MIT dissertation.
Duanmu, San. (2008) *Syllable Structure: The Limits of Variation*. Oxford: Oxford University Press.
Dudas, Karen Marie. (1976) *The Phonology and Morphology of Modern Javanese.* University of Illinois dissertation.
Duponceau, Peter. (1818) English phonology; or, an essay towards an analysis and description of the component sounds of the English language. *Transactions of the American Philosophical Society (new series)* 1, 228–264.
Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. (1999) Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25, 1568–1578.
Dupoux, Emmanuel, Christophe Pallier, and Kazuhiko Kakehi. (2001) New evidence for prelexical phonological processing in spoken word recognition. *Language and Cognitive Processes* 16, 491–505.
Durand, Jacques (ed.) (1986) *Dependency and Non-linear Phonology*. London: Croom Helm.
Durand, Jacques. (1990) In defence of dependency phonology. *Rivista di Linguistics* 2, 87–102.
Durand, Jacques and Bernard Laks. (2000) Relire les phonologues du français:

Maurice Grammont et la loi des trois consonnes. *Langue Française* 126: 29–39.
Dürr, Michael. (1987) A preliminary reconstruction of the Proto-Mixtec tonal system. *Berlin: Indiana* 11, 19–62.
Dyck, Carrie. (1995) *Constraining the Phonology-Phonetics Interface with Exemplification from Spanish and Italian Dialects*. University of Toronto dissertation.
Dye, Amanda. (2006) Phonological rule-ordering in artificial language learning. Manuscript, Harvard University.
Dye, Wayne. (1992) Bahinemo Organised Phonology Data. SIL Papua New Guinea. Retrieved April 8, 2011 from http://www.sil.org/pacific/png/pubs/0000048/Bahinemo.pdf.
Eberhard, David. (2007) Mamainde tone, in *Language Endangerment and Endangered Languages*, ed. by Leo Wetzels, 287–308, CNWS Publications.
Ebert, Karen. (1979) *Sprache und Tradition der Kera (Tschad), Teil III: Grammatik*. Berlin: Reimer.
Echeverría, Max and Helen Contreras. (1965) Araucanian phonemics. *International Journal of American Linguistics* 31, 132–135.
Eckert, Penelope. (2004) The Meaning of Style, in *Proceedings of the 11th Annual Symposium about Language and Society – Austin*, ed. by L. Mahalingappa Chiang, E. Chun, and S. Mehus.
Eddington, David. (2003) Issues in modeling language processing analogically. *Lingua* 114, 849–871.
Edmondson, Jerold A. and Kenneth J. Gregerson. (1992) On five-level tone systems, in *Language in Context*, ed. by Shina Ja J. Huang and William R. Merrifield, 555–576, SIL and University of Texas, Arlington.
Edmondson, T. and J. T. Bendor-Samuel. (1966) Tone patterns in Etung. *Journal of African Languages* 5, 1–6.
Edwards, J., Mary E. Beckman, and B. Munson. (2004) The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research* 47, 421–436.
Efere, Emmanuel Efereala. (2001) The pitch system of the Bumo dialects of Izon, in *Current Research on African Languages and Linguistics*, ed. by Suzanne Gessner, Sunyoung Oh, and Kayono Shiobara, volume 4, 115–259, UBCWPL.
Egerod, Søren. (1966) A statement on Atayal phonology, in *Artibus Asiae Supplementum XXIII (Felicitation volume for the 75th birthday of Professor G. H. Luce)*, 120–30, Ascona, Switzerland.
Eisenstat, Sarah. (2009) *Learning underlying forms with MaxEnt*. Master's thesis, Brown University.
Eisner, Jason. (1997) Efficient generation in Primitive Optimality Theory, in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, 313–320, East Stroudsburg, PA: Association for Computational Linguistics.
Eisner, Jason. (1998) FootForm decomposed: Using primitive constraints in OT, in *Proceedings of SCIL VIII*, ed. by Benjamin Bruening, number 31 in MIT *Working Papers in Linguistics*, 115–143, Cambridge, MA.
Eisner, Jason. (2001) Expectation semirings: Flexible EM for finite-state transducers, in *Proceedings of the ESSLLI Workshop on Finite-State Methods in Natural Language Processing*, ed. by Gertjan van Noord.
Eisner, Jason. (2002) Parameter estimation for probabilistic finite-state transducers, in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, p. 108, East Stroudsburg, PA: Association for Computational Linguistics.
Elenbaas, Nine. (1999) *A Unified Account of Binary and Ternary Stress*. Utrecht:

Netherlands Graduate School of Linguistics.
Elenbaas, Nine and René Kager. (1999) Ternary rhythm and the lapse constraint. *Phonology* 16, 273–329.
Elfner, Emily. (2011) The interaction of linearization and prosody: Evidence from pronoun postposing in Irish, in *Formal Approaches to Celtic Linguistics*, ed. by Andrew Carnie, Cambridge: Cambridge Scholars Publishing.
Elfner, Emily and Wendell Kimper. (2008) Reduplication without RED: Evidence from *diddly*-infixation, in *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. by Natasha Abner and Jason Bishop, volume 150–158, Somerville, MA: Cascadilla Press.
Eliason, Norman E. (1942) On syllable division in phonemics. *Language* 18(2), 144–147.
Ellington, John. (1977) *Aspects of the Tiene Language*. University of Wisconsin, Madison dissertation.
Ellis, Lucy and William J. Hardcastle. (2002) Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30, 373–396.
Ellison, T. Mark. (1994) Phonological derivation in Optimality Theory, in *Proceedings of the 15th International Conference on Computational Linguistics*, volume 2, 1007–1013.
Elman, Jeffrey. (1990) Finding structure in time. *Cognitive Science* 14, 179–211.
Elmedlaoui, Mohamad. (1992) *Aspects des représentations phonologiques dans certaines langues chamito-sémitiques*. Université Mohamad V, Rabat dissertation.
Elordieta, Gorka. (1997) Accent, tone, and intonation in Lekeitio Basque, in *Issues in the Phonology and Morphology of the Major Iberian Languages*, ed. by F. Martinez-Giland and A. Morales-Front, 4–78, Washington, DC: Georgetown University Press.
Elordieta, Gorka. (1998) Intonation in a pitch-accent dialect of Basque. *International Journal of Basque Linguistics and Philology* 32, 511–569.
Elordieta, Gorka. (2006) Binarity at the left edge of intonational phrases, in *GLOW Workshop on Prosodic Phrasing*, April 2006, Universitat Autònoma de Barcelona.
Elordieta, Gorka. (2007a) A constraint-based analysis of restrictions on intonational realization of focus in Northern Bizkaian Basque, in *Tones and Tunes I: Typological Studies in Word and Sentence Prosody*, ed. by Tomas Riad and Carlos Gussenhoven, 199–232, Berlin: Mouton de Gruyter.
Elordieta, Gorka. (2007b) Constraints on intonational prominence of focalized constituents, in *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*, ed. by Chungmin Lee, Matthew Gordon, and Daniel Buring, Dordrecht: Springer.
Elordieta, Gorka. (2007c) Minimum size constraints on intermediate phrases, in *Proceedings of the 16th International Congress of Phonetic Sciences*, 1021–1024.
Elordieta, Gorka. (2007d) Segmental phonology and syntactic structure, in *The Oxford Handbook of Linguistic Interfaces*, ed. by Gillian Ramchand and Charles Reiss, 125–178, Oxford: Oxford University Press.
Elordieta, Gorka, Sónia Frota, Pilar Prietó and Marina Vigário. (2003) Effects of constituent length and syntactic branching on intonational phrasing in Ibero-Romance, in *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. by M. J. Solé, D. Recasens, and I. Romero, 487–490, Barcelona, Universitat Autònoma de Barcelona.
Elordieta, Gorka, Sónia Frota, and Marina Vigário. (2005) Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica* 59, 110–143.

Embick, David and Rolf Noyer. (2007) Distributed morphology and the syntax-morphology interface, in *The Oxford Handbook of Linguistic Interfaces*, ed. by Gillian Ramchand and Charles Reiss, 289–324, Oxford: Oxford University Press.

Emeneau, M. B. (1961) *Kolami: A Dravidian Language*. Annamalainagar: Annamalai University.

Emmorey, Karen, Thomas Grabowski, Steven McCullough, Hannah Damasio, Laurie Ponto, Richard Hichwa and Ursula Bellugi. (2004) Motor-iconicity of sign language does not alter the neural systems underlying tool and action naming. *Brain and Language* 89, 27–38.

Emonds, Joseph. (1976) *A Transformational Approach to English Syntax*. New York: Academic Press.

Engberg-Pederson, Elisabeth. (1993) *Space in Danish Sign Language*. Hamburg: Signum Verlag.

England, Nora. (1983) *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

Epps, Patience. (2005) *A Grammar of Hup*. University of Virginia dissertation.

Epstein, Melissa Ann. (2002) *Voice Quality and Prosody in English*. UCLA dissertation.

Erickson, Donna. (1978) *A Physiological Analysis of the Tones of Thai*. University of Connecticut dissertation.

Erickson, Donna. (1993) Laryngeal muscle activity in connection with Thai tones. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, University of Tokyo* 27, 135–149.

Erickson, Donna, Kiyoshi Honda, Hiroyuki Hirai and Mary E. Beckman. (1995) The production of low tones in English intonation. *Journal of Phonetics* 23, 179–188.

Ernestus, Mirjam. (2000) *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-based Study of the Phonology-Phonetics Interface*. Utrecht: LOT.

Ernestus, Mirjam and R. Harald Baayen. (2003) Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79, 5–38.

Ernestus, Mirjam and R. Harald Baayen. (2004) Analogical effects in regular past tense production in Dutch. *Linguistics* 42, 873–903.

Ernestus, Mirjam, R. Harald Baayen, and Rob Schreuder. (2002) The recognition of reduced word forms. *Brain and Language* 81, 162–173.

Ernestus, Mirjam, Mybeth Lahey, Femke Verhees, and R. Harald Baayen. (2006) Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America* 120, 1040–1051.

Escudero, Paola. (2000) *Developmental Patterns in the Adult L2 Acquisition of New Contrasts: The Acoustic Cue Weighting in the Perception of Scottish Tense/Lax Vowels by Spanish Speakers.* Utrecht: LOT 113.

Escudero, Paola. (2005) *Linguistic Perception and Second Language Aquisition: Explaining the Attainment of Optimal Phonological Categorization*, volume 113. Utrecht: LOT.

Esper, Erwin. (1925) *A Technique for the Experimental Investigation of Associative Interference in Artificial Linguistic Material*. Linguistic Society of America.

Ettlinger, Marc. (2007) Variation as a window into opacity, in *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*, ed. by Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki, and Suwon Yoon, Chicago: Chicago Lingustic Society.

Ettlinger, Marc. (2008) *Input-driven Opacity*. University of California Berkeley dissertation.

Evans, Nicolas. (1995) Current issues in the phonology of Australian languages, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 723–761, Oxford: Blackwell.

Everett, Daniel. (1988) On metrical constituent structure in Pirahã phonology. *Natural Language and Linguistic Theory* 6, 207–246.

Everett, Daniel and Keren Everett. (1984) On the relevance of syllable onsets to stress placement. *Linguistic Inquiry* 15, 705–711.

Ewen, Colin J. (1980) *Aspects of Phonological Structure With Particular Reference to English and Dutch*. University of Edinburgh dissertation.

Ewen, Colin J. (1982) The internal structure of complex segments, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 2, Dordrecht: Foris.

Ewen, Colin J. (1986) Segmental and suprasegmental structure, in *Dependency and Non-linear Phonology*, ed. by Jacques Durand, 203–222, London: Croom Helm.

Ewen, Colin J. (1993) Dependency phonology, in *The Encyclopaedia of Language and Linguistics*, ed. by R. E. Asher, 864–867, Oxford: Pergamon Press.

Ewen, Colin J. (1995) Dependency relations in phonology, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 570–585, Oxford: Blackwell.

Eychenne, Julien. (2006) *Aspects de la phonologie du schwa dans le français contemporain: optimalité, visibilité prosodique, gradience*. Université de Toulouse – Le Mirail dissertation.

van den Eynde, Karel. (1968) *Eléments de grammaire yaka*. Université Lovanium.

Fabiano-Smith, Leah and Jessica Barlow. (2010) Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *Internatioanl Journal of Bilingual Education and Bilingualism* 13, 81–97.

Fant, Gunnar. (1960) *Acoustic Theory of Speech Production*. The Hague: Mouton.

Fasold, Ralph W. (1972) *Tense Marking in Black English*. Washington, DC: Center for Applied Linguistics.

Fee, E. J. (1995) Segments and syllables in early language acquisition, in *Phonological Acquisition and Phonological Theory*, ed. by J. Archibald, 43–61, Hillsdale, NJ: Lawrence Erlbaum Associates.

Feldman, Naomi H., Thomas L. Griffiths, and James L. Morgan. (2009) The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review* 116, 752–782.

Fenlon, Jordan, Tanya Denmark, and Bencie Woll. (2006) Seeing sentence boundaries. *Sign Language and Linguistics* 10, 2, 177–200.

Ferguson, Charles A. and Carol B. Farwell. (1975) Words and sounds in early language acquisition. *Language* 51, 419–439.

Ferguson, Charles A., Lise Menn, and C. Stoel-Gammon. (1992) *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press.

Féry, Caroline. (no date) German Sentence Accents and Embedded Prosodic Phrases. Manuscript.

Féry, Caroline and Fabian Schubö. (2008) Hierarchical prosodic structures in the intonation of syntactically recursive sentences. Manuscript University of Potsdam.

Féry, Caroline and Hubert Truckenbrodt. (2005) Sisterhood and tonal scaling. *Studia Linguistica. Special Issue: Boundaries in Intonational Phonology* 59, 223–243.

Féry, Caroline and Ruben van de Vijver. (2003) *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press.

Féry, Caroline and Vieri Samek-Lodovici. (2006) Focus projection and prosodic prominence in nested foci. *Language* 82, 131–150.

Fidelholtz, James. (1975) Word frequency and vowel reduction in English. *Papers from the 11th Annual Meeting of the Chicago Linguistic Society*, 200–213, Chicago: Chicago Linguistic Society.

Field, Frederic. (2002) *Linguistic Borrowing in Bilingual Contexts*. Amsterdam: Benjamins.

Fikkert, Paula. (1994) *On the Acquisition of Prosodic Structure*. Rijksuniversiteit te Leiden dissertation.

Finley, Sara. (2008) *Formal and Cognitive Restrictions on Vowel Harmony*. Johns Hopkins University dissertation.

Firchow, Irwin B. and Jacqueline Firchow. (1969) An abbreviated phoneme inventory. *Anthropological Linguistics* 11, 271–276.

Firth, J. R. (1948) Sounds and prosodies. *Transactions of the Philological Society* 48, 127–152.

Fischer, Markus. (2005) A Robbins-Monro type learning algorithm for an entropy maximizing version of stochastic Optimality Theory. Master's thesis, Humboldt University, Berlin.

Fischer-Jørgensen, Eli (1952) On the definition of phoneme categories on a distributional basis. *Acta Linguistica*, 8–39.

Fisher, R. A. (1922) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 222, 309–368.

Fisher, Ryan. (2004) *L'influence de la graphie hors-malformation dans les emprunts anglais en italien calabrais*. Quebec City: Project CoPho Research Report, Laval University.

Fisher, William, George Doddington, and Kathleen M. Goudie-Marshall. (1986) The DARPA speech recognition research database: Specification and status. *Proceedings of the DARPA Speech Recognition Workshop* 12, 93–100.

Fitzgerald, Colleen. (2012) Prosodic inconsistency in Tohono O'odham. *International Journal of American Linguistics* 78.

Flege, James and Wieke Eefting. (1986) Developmental and linguistic effects of English stops by native Spanish speakers. *Phonetica*, 43, 155–171.

Flege, James and Wieke Eefting. (1987) Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15, 67–83.

Flemming, Edward. (1993) *The Role of Metrical Structure in Segmental Rules*. Master's thesis, UCLA. A shorter version is available in *NELS 24* volume 1, 97–110, 1994.

Flemming, Edward. (1995) Vowels undergo consonant harmony. Presented at the Trilateral Phonology Weekend No. 5.

Flemming, Edward. (1996) Evidence for constraints on contrast: The dispersion theory of contrast. *UCLA Working Papers in Phonology* 1, 86–106.

Flemming, Edward. (2001a) *Auditory Representations in Phonology*. New York: Garland Press.

Flemming, Edward. (2001b) Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, 7–44.

Flemming, Edward. (2004) Contrast and perceptual distinctiveness, in *Phonetically Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 232–276, Cambridge: Cambridge University Press.

Flemming, Edward and Stephanie Johnson. (2007) Rosa's roses: Reduced vowels in American English. *Journal of the International Phonetic Association* 37, 83–96.

Fletcher, Janet. (2010) The prosody of speech: Timing and rhythm, in *Handbook of Phonetic Sciences*, ed. by William J. Hardcastle, J. D. M. H. Laver, and Fiona E. Gibbon, 523–602, Oxford: Wiley-Blackwell, 2nd edition.

Fodor, Jerry A. (1981) Introduction: Some notes on what linguistics is about, in *Readings in Philosophy of Psychology*, ed. by Ned Block, 197–202, Cambridge, MA: Harvard University Press.

Fodor, Jerry A. and J. J. Katz. (1964) *The Structure of Language: Readings in the Philosophy of Language.* Englewood Cliffs, NJ: Prentice-Hall.

Fodor, Jerry A. and Z. W. Pylyshyn. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3–71.

Fontaney, L. D. (1980) Le verbe, in *Eléments de description du punu*, ed. by F. Nsukua-Nkutsi, 51–114, CRLS Université Lyon II.

Ford, Alan and Rajendra Singh. (1983) On the status of morphophonology, in *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*, ed. by John F. Richardson, M. Marks, and A. Chukerman, 63–78, Chicago: Chicago Linguistic Society.

Ford, Alan and Rajendra Singh. (1985) Towards a non-parametric morphology, in *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, ed. by M. Niepokuj, M. VanClay, N. Nikiforidou, and D. Feder, 87–95, Berkeley Linguistic Society.

Ford, Kevin. (1993) A preliminary comparison of Kamano-Yagaria. *Language and Linguistics in Melanesia* 24, 191–202.

Fortesque, Michael. (1984) *West Greenlandic*. London: Croom Helm.

Foster, Michael. (1982) Alternating weak and strong syllables in Cayuga words. *International Journal of American Linguistics* 48, 59–72.

Fouché, P. (1959) *Traité de la prononciation française*. Paris: Klincksieck.

Fougeron, Cécile and Pat Keating. (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106, 3728–3740.

Fougeron, Cécile and Donca Steriade. (1997) Does the deletion of French schwa lead to the neutralization of lexical distinctions?, in *Proceedings of the Eurospeech Conference*, 943–946, Rhodes.

Fourakis, Marios and Gregory K. Iverson. (1984) On the "incomplete neutralization" of German final obstruents. *Phonetica* 41, 140–149.

Fourakis, Marios and Robert Port. (1986) Stop epenthesis in English. *Journal of Phonetics* 14, 197–221.

Fowler, Carol A. (1981) Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research* 46, 127–139.

Fox, Robert A. and Dale Terbeek. (1977) Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics* 5, 27–34.

Fox-Tree, J. E. and H. H. Clark. (1997) Pronouncing "the" as "thee" to signal problems in speaking. *Cognition* 62, 151–167.

Francis, W. Nelson and Henry Kučera. (1982) *Frequency Analysis of English Usage*. Boston: Houghton Mifflin.

Frank, Robert and Giorgio Satta. (1998) Optimality Theory and the generative complexity of constraint violability. *Computational Linguistics* 24, 307–315.

Frank, Robert and Shyam Kapur. (1996) On the use of triggers in parameter setting. *Linguistic Inquiry* 27, 623–660.

Franks, Steven. (1987) Regular and irregular stress in Macedonian. *International Journal of Slavic Linguistics and Poetics* 35, 93–142.

Fraser, N. M. (2005) Dependency grammar, in *Encyclopedia of Language and Lingusitics*, Oxford: Pergamon Press, 2nd edition.

Friederici, Angela D. and Jeanine E. Wessels. (1993) Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics* 54, 287–295.

Friedman, Lynn. (1976) *Phonology of a Soundless Language: Phonological Structure of American Sign Language*. University of California Berkeley dissertation.

Frisch, Stefan. (1996) *Similarity and Frequency in Phonology*. Northwestern University dissertation.

Frisch, Stefan, Nathan R. Large, and David B. Pisoni. (2000) Perception of wordlikeness: Effects of segment

probabiliy and length on the processing of nonwords. *Journal of Memory and Language* 42, 481–496.
Frisch, Stefan, Nathan R. Large, Bushra Zawaydeh and David B. Pisoni. (2001) Emergent phonotactic generalizations in English and Arabic, in *Frequency and the Emergence of Linguistic Structure*, ed. by Joan L. Bybee and Paul Hopper, 159–179, Amsterdam: Benjamins.
Frisch, Stefan, Janet B. Pierrehumbert, and Michael Broe. (2004) Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22, 179–228.
Frisch, Stefan and Bushra Zawaydeh. (2001) The psychological reality of OCP-place in Arabic. *Language* 77, 91–106.
Frishberg, Nancy. (1975) Arbitrariness and iconicity. *Language* 51, 696–719.
Fromkin, Victoria A. (1971) The non-anomalous nature of anomalous utterances. *Language* 41, 27–52.
Frota, Sónia. (2000) *Prosody and Focus in European Portuguese: Phonological Phrasing and Intonation*. New York, London: Garland Publishing Co.
Fry, D. B. (1968) Prosodic phenomena, in *Manual of Phonetics*, ed. by Bertil Malmerg, 364–410, Amsterdam: North-Holland.
Fudge, Erik. (1969) Syllables. *Journal of Linguistics* 3, 253–286.
Fudge, Erik. (1987) Branching structure within the syllable. *Journal of Linguistics* 23, 359–377.
Fujimura, Osamu. (1992) Phonology and phonetics – a syllable-based model of articulatory organization. *Journal of the Acoustical Society of Japan* 13, 39–48.
Fujisaki, Hiroya. (1983) Dynamic characteristics of voice fundamental frequency in speech and singing, in *The Production of Speech*, ed. by Peter F. MacNeilage, 39–55, New York: Springer-Verlag.
Fujisaki, Hiroya and H. Sudo. (1971) Synthesis by rule of prosodic features of connected Japanese, in *Proceedings of the 7th International Congress on Acoustics (ICA), Budapest*, 133–136.
Fulmer, S. Lee. (1991) Dual-position affixes in Afar: An argument for phonologically-driven morphology, in *Proceedings of the Ninth Annual Meeting of the West Coast Conference on Formal Linguistics*, ed. by A. Halpern, 189–204, Stanford, CA: CSLI Publications.
Fung, Roxana Suk-Yee. (2000) *Final Particles in Standard Cantonese: Semantic Extension and Pragmatic Inference*. Ohio State University dissertation.
Furby, Christine. (1974) *Garawa Phonology*. Canberra: Australian National University.
Furere, Rutinywa and Annie Rialland. (1983) Accent tonal en kinyarwanda, in *Current Approaches to African Linguistics*, ed. by Jonathan Kaye *et al.*, volume 2, 139–150, Dordrecht: Foris.
Gabas, Nilson Jr. (1999) *A Grammar of Karo, Tupí (Brazil)*. University of California, Santa Barbara dissertation.
Gafos, Adamantios. (1996) *The Articulatory Basis of Locality in Phonology.* Johns Hopkins University dissertation. Published by Garland Publishing, New York, 1999.
Gafos, Adamantios. (1998a) A-templatic reduplication. *Linguistic Inquiry* 515–525.
Gafos, Adamantios. (1998b) Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16, 223–78.
Gafos, Adamantios. (2002) A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269–337.
Gafos, Adamantios and Stefan Benus. (2006) Dynamics of phonological cognition. *Cognitive Science* 30, 1–39.
Gafos, Adamantios and Linda Lombardi. (1999) Consonant transparency and vowel echo. *Proceedings of the 29th Annual Meeting of the North East Linguistic Society*, 81–96.
Gagné, Christina. (2001) Relation and lexical priming during the interpretation

of noun-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27, 236–254.
Gahl, S. (2008) "Time" and "thyme" are not homophones: The effect of lemma frequency and word durations in a corpus of spontaneous speech. *Language* 84, 474–496.
Gairdner, W. H. T. (1925) *The Phonetics of Arabic.* London: Oxford University Press.
Gårding, Eva. (1977) The importance of turning points for the pitch patterns of Swedish accents, in *Studies in Stress and Accent*, ed. by Larry M. Hyman, volume 4, 27–35, Los Angeles: University of Southern California.
Gårding, Eva. (1993) On parameters and principles in intonation analysis, in *Working Papers, Lund University, Department of Linguistics*, volume 40, 25–47, Lund University.
Gårding, Eva, Osamu Fujimura, and Hajime Hirose. (1970) Laryngeal control of Swedish word tones: A preliminary report on an EMG study. Technical report, Annual Report of the Research Institute on Logopedics and Phoniatrics 4, 45–54.
Gårding, Eva, Paul Kratochvil, Jan-Olaf Svantesson, and Jialu Zhang. (1986) Tone 3 and tone 4 identification in Modern Standard Chinese. *Language and Speech* 29, 281–293.
Garrett, Andrew. (1990) The origin of NP split ergativity. *Language* 66, 261–296.
Garrett, Andrew and Juliette Blevins. (2009) Analogical morphophonology, in *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 527–546, Cambridge, MA: MIT Press.
Gaskell, M. Gareth. (2003) Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics* 31, 447–463.
Gaskell, M. Gareth and William D. Marslen-Wilson. (1996) Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 22, 144–158.
Gaskell, M. Gareth and William D. Marslen-Wilson. (1997) Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes* 12, 613–656.
Gaskell, M. Gareth, Philip T. Quinlan, Jakke Tamminen, and Alexandra A. Cleland. (2008) The nature of phoneme representations in spoken word recognition. *Journal of Experimental Psychology: General* 137, 282–302.
Gathercole, Susan E., Clive R. Frankish, Susan J. Pickering and Sarah Peaker. (1999) Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25, 84–95.
Gathercole, Susan E. and Amanda J. Martin. (1996) Interaction processes in phonological memory, in *Models of Short Term Memory*, ed. by Susan E. Gathercole, 73–100, Hove, UK: Psychology Press.
Gathercole, Susan, C. S. Willis, and Alan D. Baddeley (1991) The influences of number of syllables and word-likeness on children's repetition of nonwords. *Applied Psycholinguistics* 12, 349–367.
Gazdar, Gerald. (1982) Phrase structure grammar, in *The Nature of Syntactic Representations*, ed. by Pauline Jacobsen and Geoff K. Pullum, 131–186, Dordrecht: D. Reidel.
Gazdar, Gerald, E. Klein, Geoffrey Pullum, and Ivan Sag. (1985) *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
Gendron, Jean Denis. (1966) *Tendances phonétiques français parlé au Canada*. Quebec, P.Q.: Les Presses de L'Université Laval.
Geraci, Carlo. (2009) Epenthesis in Italian Sign Language. *Sign Language Linguistics* 12, 1, 3–51.
Gerfen, Chip. (2002) Andalusian codas. *Probus* 14, 247–277.
Gerfen, Chip and Kathleen Currie Hall, Coda aspiration and incomplete

neutralisation in Eastern Andalusian Spanish. Manuscript.
Gerken, LouAnn. (1994) A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language* 21, 565–584.
Gerken, LouAnn. (1996) Prosodic structure in young children's language production. *Language* 72, 683–712.
Gerken, LouAnn and Alex Bollt. (2008) Three exemplars allow at least some linguistic generalizations: implications for generalization mechanisms and constraints. *Language Learning and Development* 4, 228–248.
Gerken, LouAnn and B. J. McIntosh. (1993) Interplay of function morphemes and prosody in early language. *Developmental Psychology* 29, 448–457.
Ghazeli, Salem. (1977) *Back Consonants and Backing Coarticulation in Arabic.* University of Texas, Austin dissertation.
Ghini, Mirco. (1993) *ϕ*-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12, 41–78.
Giannini, Antonella and Umberto Cinque. (1978) *Phonetic Status and Phonemic Function of the Final Devoiced Stops in Polish*. Naples: Seminario di studi dell'Europa orientale.
Gibbons, A. (1985) *Algorithmic Graph Theory*. Cambridge: Cambridge University Press.
Gibson, Edward and Kenneth Wexler. (1994) Triggers. *Linguistic Inquiry* 25, 407–454.
Gibson, Lorna F. (1956) Pame (Otomi) phonemics and morphophonemics. *International Journal of American Linguistics* 22, 242–265.
Gick, B. (2007) A lingual motor differentiation model for liquid substitutions in children's speech. Presented at the ASA meeting.
Gick, Bryan, Douglas Pulleyblank, Fiona Campbell and Ngessimo Mutaka. (2006) Low vowels and transparency in Kinande vowel harmony. *Phonology* 23, 1–20.
Gierut, J. A. (1996) Categorization and feature specification in phonological acquisition. *Journal of Child Language* 23, 397–415.
Gierut, J. A. (1999) Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research* 42, 708–726.
Gil, David. (2002) Ludlings in Malayic Languages: An Introduction, in *PELBBA 15, Pertemuan Linguistik (Pusat Kajian) Bahasa, dan Budaya Atma Jaya: Kelima Belas*, Jakarta: Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya.
Gildea, Daniel and Daniel Jurafsky. (1996) Learning bias and phonological rule induction. *Computational Linguistics* 22, 497–530.
Gilley, Leoma G. (1992) *An Autosegmental Approach to Shilluk Phonology*. SIL and University of Texas at Arlington.
Gimson, A. C. (1970) *An Introduction to the Pronunciation of English*. London: Edward Arnold.
Ginneken, Jacques Van. (1907) *Principes de linguistique psychologique: essai de synthèse*. Paris: Marcel Rivière.
Girón Higuita, Jesús María and W. Leo Wetzels. (2007) Tone in Wãnsöhöt (Puinave), in *Language Endangerment and Endangered Languages*, ed. by Leo Wetzels, 129–156, Leiden: CNWS Publications.
Gleason, Jean Berko and Nan Bernstein Ratner (eds.) (1998) *Psycholinguistics*. Belmont, CA: Wadsworth/Thomson Learning.
Glover, Bonnie. (1988) *The Morphophonology of Muscat Arabic*. UCLA dissertation.
Gnanadesikan, Amalia. (1997) *Ternary Scales in Phonology*. University of Massachusetts Amherst dissertation.
Gnanadesikan, Amalia. (2004) Markedness and faithfulness constraints in child phonology, in *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, Cambridge: Cambridge University Press.

Goad, Heather. (1993) *On the Configuration of Height Features*. University of Southern California dissertation.
Goad, Heather. (1997) Consonant harmony in child language: An Optimality-Theoretic account, in *Focus on Phonological Acquisition*, ed. by S. J. Hannahs and M. Young-Scholten, 113–142, Amsterdam: Benjamins.
Goad, Heather and K. Brannen. (2003) Phonetic evidence for phonological structure in syllabification, in *The Phonological Spectrum: Suprasegmental Structure*, ed. by Jeroen van de Weijer, Vincent J. van Heuven, and Harry van der Hulst, 3–30, Amsterdam: Benjamins.
Goad, Heather and M. Buckley. (2006) Prosodic structure in child French: Evidence for the foot. *Catalan Journal of Linguistics* 5, 109–142.
Goad, Heather and Y. Rose. (2003) Segmental-prosodic interaction in phonological development: A comparative investigation. *Canadian Journal of Linguistics* 48, 139–152.
Goad, Heather and Y. Rose. (2004) Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic, in *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 109–157, Cambridge University Press.
Godfrey, John J., Edward C. Holliman, and Jane McDaniel. (1992) SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* 517–520, San Francisco.
Goedemans, Rob. (1996) An optimality account on onset-sensitive stress in quantity-insensitive languages. *Linguistic Review* 13, 33–47.
Goedemans, Rob. (1998) *Weightless Segments*. The Hague: Holland Academic Graphics.
Goeman, A. C. M. (1999) *T-deletie in Nederlandse Dialecten. Kwantitatieve Analyse van Structurele, Ruimtelijke en Temporele Varietie*. The Hague: Vrije Universiteit, Amsterdam dissertation.
Goeman, A. C. M. and Johan Taeldeman. (1996) Fonologie en Morfologie van de Nederlandse Dialecten. Een Nieuwe Materiaal-Verzameling en Twee Nieuwe Atlasprojecten. *Taal en Tongval* 48, 38–59.
Goeman, A. C. M. and P. Th. van Reenen. (1985) Word-Final *t*-Deletion in Dutch dialects: The roles of conceptual prominence, articulatory complexity, paradigmatic properties, token frequency and geographical distribution, in *Vrije Universiteit Working Papers in Linguistics*, volume 16, Amsterdam: Vakgroep Algemene Taalwetenschap Vrije Universiteit.
Gold, E. Mark. (1967) Language identification in the limit. *Information and Control* 10(5), 447–474.
Goldin-Meadow, Susan. (2003) *The Resilience of Language*. New York: Psychology Press.
Goldinger, Stephen D. (1996) Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 1166–1182.
Goldinger, Stephen D. (1997) Word and voices: Perception and production in an episodic lexicon, in *Talker Variability in Speech Processing*, ed. by K. Johnson and J. W. Mullenix, 33–66, San Diego, CA: Academic Press.
Goldinger, Stephen D. (1998) Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–279.
Goldinger, Stephen D. (2007) A complementary-systems approach to abstract and episodic speech perception, in *Proceedings of the 16th International Congress of Phonetic Sciences*, 49–54, Saarbrücken.
Goldinger, Stephen D. and Tamiko Azuma. (2004) Episodic memory reflected in

printed word naming. *Psychonomic Bulletin and Review* 11, 716–722.
Goldrick, Matthew. (2001) Turbid output representations and the unity of opacity, in *Proceedings of the Northeast Linguistics Society 30*, ed. by M. Hirotani, A. Coetzee, N. Hall, and J.-Y. Kim, volume 1, 231–245, Amherst, MA: GLSA.
Goldrick, Matthew. (2004) Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language* 51, 586–603.
Goldrick, Matthew. (2006) Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes* 21, 817–855.
Goldrick, Matthew. (2008) A harmonic grammar account of lexically-conditioned phonetic variation. Presented at the Linguistic Society of America Annual Meeting.
Goldrick, Matthew and Robert Daland. (2009) Linking speech errors and phonological grammars: Insights from harmonic grammar networks. *Phonology* 26, 147–185.
Goldrick, Matthew and Meredith Larson. (2008) Phonotactic probability influences speech production. *Cognition* 107, 1155–1164.
Goldrick, Matthew and Meredith Larson. (2010) Constraints on the acquisition of variation, in *Papers in Laboratory Phonology X: Variation, Detail and Representation*, ed. by C. Fougeron, B. Kuhnert, M. D'Imperio, and N. Vallee, 285–310, Berlin: Mouton de Gruyter.
Goldrick, Matthew and Brenda Rapp. (2007) Lexical and post-lexical phonological representations in spoken production. *Cognition* 102, 219–260.
Goldsmith, John A. (1976a) *Autosegmental Phonology*. MIT dissertation. Published by Garland Publishing, New York, 1979.
Goldsmith, John A. (1976b) An overview of autosegmental phonology. *Linguistic Analysis* 2, 23–68.
Goldsmith, John A. (1978) English as a tone language. *Communication and Cognition* 11, 453–476.
Goldsmith, John A. (1984) Tone and accent in Tonga, in *Autosegmental Studies in Bantu Tone*, ed. by G. N. Clements and John A. Goldsmith, 19–51, Dordrecht: Foris.
Goldsmith, John A. (1985) Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2, 253–275.
Goldsmith, John A. (1987) Tone and accent, and getting the two together, in *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society*, ed. by Jon Aske, Natasha Beery, Laura Michaelis, and Hana Filip, 88–104, Berkeley Linguistics Society.
Goldsmith, John A. (1989) Autosegmental licensing, inalterability, and harmonic application, in *Papers from the 25th Annual Meeting of the Chicago Linguistic Society*, 145–156, Chicago: Chicago Linguistics Society.
Goldsmith, John A. (1990) *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
Goldsmith, John A. (1993a) Harmonic phonology, in *The Last Phonological Rule: Reflections on Constraints and Derivations*, ed. by John A. Goldsmith, 21–60, Chicago: University of Chicago Press.
Goldsmith, John A. (1993b) Introduction, in *The Last Phonological Rule*, ed. by John A. Goldsmith, 1–20, Chicago: University of Chicago Press.
Goldsmith, John A. (ed.) (1993c) *The Last Phonological Rule*. Chicago: University of Chicago Press.
Goldsmith, John A. (1995) Phonological theory, in *Handbook of Phonological Theory*, ed. by John A. Goldsmith, 1–23, Oxford: Blackwell.
Goldsmith, John A. (2001) Unsupervised learning of the morphology of a natural language. *Computational Linguistics* 27, 153–198.
Goldsmith, John A. (2006) An algorithm for the unsupervised learning of

morphology. *Natural Language Engineering* 12, 353–371.
Goldsmith, John A. (2007) Probability for linguists. *Mathématiques et Sciences Humaines* 180(4), 73–98.
Goldsmith, John A. (2011) Segmentation and morphology, in *The Handbook of Computational Linguistics and Natural Language Processing*, ed. by Alex Clark, Chris Fox, and Shalom Lappin, 364–393, Oxford: Wiley-Blackwell.
Goldsmith, John A. and Gary Larson. (1990) Local modeling and syllabification, in *Papers from the 26th Annual Meeting of the Chicago Linguistic Society*, 129–142, Chicago: Chicago Linguistic Society.
Goldsmith, John A. and Gary Larson. (1993) Using networks in a harmonic phonology, in *Papers from the 28th Annual Meeting of the Chicago Linguistic Society*, ed. by Costas Canakis, Grace Chan, and Jeannette Denton, volume 2, 94–125, Chicago: Chicago Linguistic Society.
Goldsmith, John A. and Jason Riggle. (2012) Information theoretic approaches to phonology: The case of Finnish vowel harmony. *Natural Language and Linguistic Theory*.
Goldsmith, John A. and Firmard Sabimana. (1986) The Kirundi verb, in *Modèles en tonologie*, ed. by Francis Jouannet, 19–62, Paris: Editions du CNRS.
Goldsmith, John A. and Aris Xanthos. (2009) Learning phonological categories. *Language* 85, 4–38.
Goldstein, Louis. (1989) On the domain of quantal theory. *Journal of Phonetics* 17, 91–97.
Goldstein, Louis and Carol A. Fowler. (2003) Articulatory phonology: A phonology for public use, in *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, ed. by Niels O. Schiller and Antje S. Meyer, 159–207, Berlin: Mouton de Gruyter.
Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman and Dani Byrd. (2007) Dynamic action units slip in speech production errors. *Cognition* 103, 386–412.
Goldwater, Sharon. (2007) *Nonparametric Bayesian Models of Lexical Acquisition*. Brown University dissertation.
Goldwater, Sharon and Mark Johnson. (2003) Learning OT constraint rankings using a maximum entropy model, in *Proceedings of the Workshop on Variation within Optimality Theory*, ed. by Jennifer Spenader, Anders Eriksson, and Osten Dahl, 111–120, Stockholm: Stockholm University.
Goldwater, Sharon and Mark Johnson. (2005) Representational bias in unsupervised learning of syllable structure, in *Proceedings of the Workshop on Psychocomputational Models of Human Language Acquisition*, ACL.
Golston, Chris and Phong Yang. (2001) White Hmong loanword phonology, in *Holland Institute for Generative Linguistics Phonology Conference V (HILP 5)*, Potsdam, Germany.
Gomez, Rebecca and LouAnn Gerken. (2000) Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* 4, 178–186.
Gomez-Imbert, Elsa. (1980) La faille tonale en tatuyo, in *Communications présentées aux Journées de Tonologie LP*, 3–121, du CNRS.
Gomez-Imbert, Elsa and Michael Kenstowicz. (2000) Barasana tone and accent. *International Journal of American Linguistics* 66, 419–463.
Goodman, Beverley (1997) Ponapean weight and [±consonantal]. Presented at the 28th Annual Meeting of the North East Linguistic Society.
Gordon, Elizabeth, Margaret Maclagan, and Jennifer Hay. (2007) The ONZE Corpus, in *Models and Methods in Handling of Unconventional Digital Corpora*, ed. by Joan C. Beal, Karen P. Corrigan, and Hermann Moisl, volume 2, 82–104, Basingstoke, UK: Palgrave.

Gordon, Matthew. (1999a) The "neutral" vowels of Finnish: How neutral are they? *Linguistica Uralica* 35, 17–21.
Gordon, Matthew. (1999b) *Syllable Weight: Phonetics, Phonology, Typology*. UCLA dissertation. Published by Routledge, New York, 2006.
Gordon, Matthew. (2000) The tonal basis of final weight criteria, in *Papers from the 36th Annual Meeting of the Chicago Linguistic Society*, 141–156, Chicago: Chicago Linguistic Society.
Gordon, Matthew. (2001) A typology of contour tone restrictions. *Studies in Language* 25, 405–444.
Gordon, Matthew. (2002a) A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20, 491–552.
Gordon, Matthew. (2002b) A phonetically driven account of syllable weight. *Language* 78, 51–80.
Gordon, Matthew. (2003) The phonology of pitch accent placement in Chickasaw. *Phonology* 20, 173–218.
Gordon, Matthew. (2004a) A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70, 1–32.
Gordon, Matthew. (2004b) Syllable weight, in *Phonetically Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 277–312, Cambridge: Cambridge University Press.
Gordon, Matthew. (2005) A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23, 595–653.
Goudbeek, Martijn. (2006) *The Acquisition of Auditory Categories*. Radboud University, Nijmegen dissertation.
Goudbeek, Martijn, Anne Cutler, and Roel Smits. (2008) Supervised and unsupervised learning of multi-dimensionally varying non-native speech categories. *Speech Communication* 50, 109–125.
Goulet, Marie-Josée. (2001) L'adaptation phonologique d'emprunts lexicaux anglais en japonais et la faible influence de la graphie. Master's thesis, Laval University, Quebec City.
Gow, D. W. (2001) Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 45, 133–159.
Grabe, Esther. (2004) Intonational variation in urban dialects of English spoken in the British Isles, in *Regional Variation in Intonation*, ed. by Peter Gilles and Jörg Peters, 9–31, Tübingen: Niemeyer.
Grammont, Maurice. (1933) *Traité de phonétique*. Paris: Delagrave.
Green, David M. (1971) Temporal auditory acuity. *Psychological Review* 78, 540–551.
Greenberg, Joseph H. (1957) *Essays in Linguistics*. Chicago: University of Chicago Press.
Greenberg, Joseph H. and James J. Jenkins. (1964) Studies in the psychological correlates of the sound system of American English. *Word* 20, 157–177.
Grice, Martine, D. Robert Ladd, and Amalia Arvaniti. (2000) On the place of phrase accents in intonational phonology. *Phonology* 17, 143–185.
Grice, Martine and Michelina Savino. (1995) Low tone versus "sag" in Bari Italian intonation: A perceptual experiment, in *Proceedings of the 13th International Congress of Phonetic Sciences*, volume 4, 658–661.
Grønnum, Nina. (1990) Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 47, 182–214.
Grønnum, Nina. (1995) Review of Möbius (1993). *Zeitschrift für Dialektologie und Linguistik* 62, 346–348.
Groce, Nora Ellen. (1985) *Everyone Here Spoke Sign Language*. Cambridge, MA: Harvard University Press.
Grose, Donovan, Ronnie Wilbur, and Katharina Schalber. (2007) Events and telicity in classifier predicates: A re-analysis of body part classifier predicates in ASL. *Lingua* 117, 1258–1284.

Grosjean, François. (1982) *Life with Two Languages: An Introduction to Bilingualism.* Cambridge, MA: Harvard University Press.

Grossberg, Stephen and Gregory Stone. (1986) Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review* 93, 46–74.

Grover, C., J. Carroll and E. Briscoe. (1993) The Alvey Natural Language Tools grammar (4th release). Technical report, Computer Laboratory, Cambridge University, UK, Technical Report 284. Retrieved April 8, 2011 from http://www.cl.cam.ac.uk/users/ejb/anlt-gram.pdf.

Grünwald, Peter. (2005) Introducing the Minimum Description Length principle, in *Advances in Minimum Description Length: Theory and Applications*, ed. by Peter Grunwald, Jae Myung, and Mark Pitt, Cambridge, MA: MIT Press.

Guerra, Anne-Marie Currie. (1999) *A Mexican Sign Language Lexicon: Internal and Cross-linguistic Similarities and Variations*. University of Texas dissertation.

Guerra, Anne-Marie Currie, Richard Meier, and Keith Walters. (2002) A crosslinguistic examination of the lexicons of four signed languages, in *Modality and Structure in Signed and Spoken Languages*, ed. by Richard Meier, K. Cormier, and D. Quinto-Pozos, 224–236, New York: Cambridge University Press.

Guerssel, Mohamed. (1978) A condition on assimilation rules. *Linguistic Analysis* 4, 225–254.

Guerssel, Mohamed. (1979) *Toward a Lexical Approach to Phonological Description*. University of Washington dissertation.

Guest, Daniel, Gary Dell, and Jennifer Cole. (2000) Violable constraints in language production: Testing the transitivity assumption of Optimality Theory. *Journal of Memory and Language* 42, 272–299.

Guimarães, Maximiliano and Andrew Nevins. (2006) Opaque nasalization in the Língua do Pe of Salvador, Brasil, in *GLOW Newsletter*.

Guion, Susan G. (1998) The role of perception in the sound change of velar palatalization. *Phonetica* 55, 18–52.

Guion, Susan G., James Flege, Reiko Akhane-Yamada and Jessica Pruitt. (2000) An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107, 2711–2724.

Guirón Higuita, J. M. and Leo Wetzels. (2007) Tone in Wãsöhöt (Puinave), in *Language Endangerment and Endangered Languages*, ed. by Leo Wetzels, 129–156, CNWS Publications.

Gupta, Abha. (1987) Hindi word stress and the obligatory branching parameter, in *Papers from the 23th Annual Meeting of the Chicago Linguistic Society*, volume 2, 134–148, Chicago: Chicago Linguistic Society.

Gupta, P. and D. S. Touretzky. (1994) Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science* 18, 1–15.

Gussenhoven, Carlos. (1984) *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.

Gussenhoven, Carlos. (1999) Discreteness and gradience in intonational contrasts. *Language and Speech* 42, 283–305.

Gussenhoven, Carlos. (2000) The lexical tone contrast of Roermond Dutch in Optimality Theory, in *Prosody, Theory and Experiment: Studies Presented to Gösta Bruce*, ed. by Merle Horne, Boston: Kluwer.

Gussenhoven, Carlos. (2004) *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. (2006) Between stress and tone in Nubi word prosody. *Phonology* 23, 193–223.

Gussenhoven, Carlos. (2007) Intonation, in *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 253–280, Cambridge: Cambridge University Press.

Gussenhoven, Carlos and Gösta Bruce. (1999) Word prosody and intonation, in *Word Prosodic Systems in the Languages of Europe*, ed. by Harry van der Hulst, 233–271, Berlin: Mouton de Gruyter.

Gussenhoven, Carlos and Toni Rietveld. (2000) The behavior of H* and L* under variations in pitch range in Dutch rising contours. *Language and Speech* 43, 183–203.

Gussmann, Edmund and Jonathan Kaye. (1993) Polish notes from a Dubrovnik café. I: The yers. *SOAS Working Papers in Linguistics and Phonetics* 3, 427–62.

Guy, Gregory R. (1980) Variation in the group and the individual: The case of final stop deletion, in *Locating Language in Time and Space*, ed. by William Labov, 1–36, New York: Academic Press.

Guy, Gregory R. (1991a) Contextual conditioning in variable lexical phonology *Language Variation and Change* 3, 223–239.

Guy, Gregory R. (1991b) Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3, 1–22.

Guy, Gregory R. (1993) The quantative analysis of linguistic variation, in *American Dialect Research. Celebrating the 100th Anniversary of the American Dialect Society, 1889–1989*, ed. by Dennis R. Preston, 223–250, Amsterdam: John Benjamins.

Guy, Gregory R. (1994) The phonology of variation, in *Papers from the 30th Annual Meeting of the Chicago Linguistic Society*, ed. by Katherine Beals *et al.*, volume 2: *The Parasession on Variation in Linguistic Theory*, 133–149, Chicago: Chicago Linguistic Society.

Guy, Gregory R. (1997) Competence, performance, and the generative grammar of variation, in *Variation, Change, and Phonological Theory*, ed. by Frans Hinskens, Roeland van Hout, and Leo Wetzels, 125–143, Amsterdam: Benjamins.

Guy, Gregory R. and Charles Boberg. (1997) Inherent variability and the Obligatory Contour Principle. *Language Variation and Change* 9, 149–164.

Haas, Mary R. (1977) Nasals and nasalization in Creek, in *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, 194–203, Berkeley, CA: Berkeley Linguistics Society.

Haiman, John. (1979) Hua: A Papuan language of New Guinea, in *Languages and Their Status*, ed. by Timothy Shopen, 35–90, Cambridge, MA: Winthrop Publishers.

Haiman, John. (1980a) Hua: A Papuan language of the Eastern Highlands of New Guinea, in *Studies in Language Companion Series*, volume 5, Amsterdam: Benjamins.

Haiman, John. (1980b) The iconicity of grammar: Isomorphism and motivation. *Language* 56, 515–540.

Hale, Kenneth. (1973) Deep-surface canonical disparities in relation to analysis and change: An Australian example. *Current Trends in Linguistics* 11, 401–458.

Hale, Kenneth and Elisabeth Selkirk. (1987) Government and tonal phrasing in Papago. *Phonology Yearbook* 4, 151–183.

Hale, Mark, Madelyn Kissock, and Charles Reiss. (2007) Microvariation, variation, and the features of Universal Grammar. *Lingua* 117, 645–665.

Hale, Mark and Charles Reiss. (1998) Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29, 656–683.

Hale, Mark and Charles Reiss. (2000a) Phonology as cognition, in *Phonological Knowledge: Conceptual and Empirical Issues*, ed. by Noel Burton-Roberts, Philip Carr, and Gerard Docherty, 161–184, Oxford: Oxford University Press.

Hale, Mark and Charles Reiss. (2000b) "Substance abuse" and

"dysfunctionalism": Current trends in phonology. *Linguistic Inquiry* 31, 157–169.
Hale, Mark and Charles Reiss. (2003) The Subset Principle in phonology: Why the tabula can't be rasa. *Journal of Linguistics* 39, 219–244.
Hale, Mark and Charles Reiss. (2008) *The Phonological Enterprise.* Oxford: Oxford University Press.
Hall, Beatrice L., R. M. R. Hall, Martin D. Pam, Amy Myers, Stephen A. Antell, and Godfrey K. Cherono. (1974) African vowel harmony systems from the vantage point of Kalenjin, *Afrika und Übersee* 57, 241–267.
Hall, Beatrice L. and R. M. R. Hall. (1980) Nez Perce vowel harmony: An Africanist explanation and some theoretical questions, in *Issues in Vowel Harmony*, ed. by Robert M. Vago, 201–236, Amsterdam: Benjamins.
Hall, Robert A. (1953) Haitian Creole: Grammar, texts, vocabulary, in *Memoirs of the American Anthropological Association*, 64, American Anthropological Association.
Hall, Tracy. (2002) Against extrasyllabic consonants in German and English. *Phonology* 19, 33–75.
Halle, Morris. (1959a) Questions of linguistics. *Il Nuovo Cimento* Supplement to Volume XIII, 494–517.
Halle, Morris. (1959b) *The Sound Pattern of Russian*. Mouton: The Hague.
Halle, Morris. (1962) Phonology in generative grammar. *Word* 18, 54–72.
Halle, Morris. (1964) On the bases of phonology, in *The Structure of Language*, ed. by Jerry A. Fodor and Jerrold J. Katz, 604–612, Englewood Cliffs, NJ: Prentice-Hall.
Halle, Morris. (1975) Confessio grammatici. *Language* 51, 525–535.
Halle, Morris. (1978) Knowledge unlearned and untaught: What speakers know about the sounds of their language, in *Linguistic Theory and Psychological Reality*, ed. by Morris Halle, Joan Bresnan, and George Miller, 294–303, Cambridge, MA: MIT Press.
Halle, Morris. (1985) Speculations about the representation of words in memory, in *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. by Victoria A. Fromkin, 101–114, San Diego, CA: Academic Press.
Halle, Morris. (1995) Feature geometry and feature spreading. *Linguistic Inquiry* 26, 1–46.
Halle, Morris and G. N. Clements. (1983) *Problem, Book in Phonology.* Cambridge, MA: MIT Press.
Halle, Morris and William Idsardi. (1995) Stress and metrical structure, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 403–443, Oxford: Blackwell.
Halle, Morris and Samuel Jay Keyser. (1971) *English Stress: Its Form, its Growth, and its Role in Verse*. New York: Harper and Row.
Halle, Morris and K. P. Mohanan. (1985) Segmental phonology of modern English. *Linguistic Inquiry* 16, 57–116.
Halle, Morris and Kenneth N. Stevens. (1959) Analysis by synthesis, in *Proceedings of the Seminar on Speech Compression and Processing, December 1959, AFCRC-TR–59–198*, ed. by L. E. Woods and W. Wathen-Dunn, volume II, paper D7.
Halle, Morris and Kenneth N. Stevens. (1971) A note on laryngeal features. *Quarterly Progress Report MIT Research Laboratory of Electronics* 101, 198–213.
Halle, Morris, Bert Vaux, and Andrew Wolfe. (2000) On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31, 387–444.
Halle, Morris and Jean-Roger Vergnaud. (1980) Three dimensional phonology. *Journal of Linguistic Research* 1, 83–105.
Halle, Morris and Jean-Roger Vergnaud. (1981) Harmony processes, in *Crossing the Boundaries in Linguistics*, ed. by

Wolfgang Klein and Willem Levelt, 1–22, Dordrecht: D. Reidel.
Halle, Morris and Jean-Roger Vergnaud. (1982) On the framework of autosegmental phonology, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, 65–82, Dordrecht: Foris.
Halle, Morris and Jean-Roger Vergnaud. (1987) *An Essay on Stress*. Cambridge, MA: MIT Press.
Hallé, Pierre. (1994) Evidence for tone-specific activity of the sternohyoid muscle in Modern Standard Chinese. *Language and Speech* 37, 103–123.
Hallé, Pierre A., Juan Segui, Uli Frauenfelder, and Christine Meunier. (1998) Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24, 592–608.
Halliday, M. A. K. (1967) *Intonation and Grammar in British English*. The Hague: Mouton.
Halliday, M. A. K. (1992) A systemic interpretation of Peking syllable finals, in *Studies in Systemic Phonology*, ed. by P. Tench, 98–121, London: Pinter.
Halpern, Aaron. (1992) *Topics in the Syntax and Placement of Clitics*. Stanford University dissertation.
Ham, William. (1998) *Phonetic and Phonological Aspects of Geminate Timing*. Cornell University dissertation. Published by Routledge, New York, 2001.
Hamann, Silke. (2003) *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT.
Hammond, Michael. (1990) The "Name Game" and onset simplification. *Phonology* 7, 159–162.
Hammond, Michael. (1997) Vowel quality and syllabification in English. *Language* 73, 1–17.
Hamp, Eric. (1976) Palatalization and harmony in Gagauz and Karaite, in *Tractata altaica: Denis Sinor, sexagenario optime de rebus altaicis merito dedicata*, ed. by W. Heissig *et al.*, 211–213, Wiesbaden: Harrassowitz.
Han, Eunjoo. (1994) *Prosodic Structure in Compounds*. Stanford University dissertation.
Hankamer, Jorge. (1973) Unacceptable ambiguity. *Linguistic Inquiry* 4, 17–68.
Hankamer, Jorge and Judith Aissen. (1974) The sonority hierarchy, in *Papers from the 20th Annual Meeting of the Chicago Linguistic Society*, 131–145, Chicago: Chicago Linguistic Society.
Hansen, Kenneth and L. E. Hansen. (1969) Pintupi phonology. *Oceanic Linguistics* 8, 153–170.
Hansen, Kenneth and L. E. Hansen. (1978) *The Core of Pintupi Grammar*. Alice Springs: Institute for Aboriginal Development, Chicago: Chicago Linguistic Society.
Hansson, Gunnar Ó. (2001a) The phonologization of production constraints: Evidence from consonant harmony, in *Papers from the 37th Annual Meeting of the Chicago Linguistic Society*, ed. by Mary Andronis, Chris Ball, Heidi Elston, and Sylvain Neuvel, volume 1, 187–200.
Hansson, Gunnar Ó. (2001b) *Theoretical and Typological Issues in Consonant Harmony*. University of California, Berkeley dissertation.
Hansson, Gunnar Ó. (2003) Laryngeal licensing and laryngeal neutralization in Faroese and Icelandic. *Nordic Journal of Linguistics* 26, 45–79.
Hansson, Gunnar Ó. (2004a) Long-distance voicing agreement: An evolutionary perspective, in *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*, ed. by Mark Ettlinger, Nicholas Fleischer, and Mischa Park-Doob, 130–141, Berkeley, Berkeley Linguistics Society.
Hansson, Gunnar Ó. (2004b) Tone and voicing agreement in Yabem, in *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, ed. by B. Schmeiser, A. Kelleher, V. Chand,

and A. Rodriguez, 318–331, Somerville, MA: Cascadilla Press.
Hansson, Gunnar Ó. (2007a) Blocking effects in agreement by correspondence. *Linguistic Inquiry* 38, 395–409.
Hansson, Gunnar Ó. (2007b) On the evolution of consonant harmony: The case of secondary articulation agreement. *Phonology* 24, 77–120.
Hansson, Gunnar Ó. (2010) *Consonant Harmony: Long-Distance Interaction in Phonology*. Berkeley, CA: University of California Press.
Hansson, Gunnar Ó. and Ronald Sprouse. (1999) Factors of change: Yowlumne vowel harmony then and now, in *Proceedings of WSCLA IV*, ed. by M. Caldecott, S. Gessner, and E.-S. Kim, 39–57, Vancouver: Department of Linguistics, University of British Columbia.
Hara, Daisuke. (2003) *A Complexity-based Approach to the Syllable Formation in Sign Language*. University of Chicago dissertation.
Haraguchi, Shosuke. (1975) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Cambridge, MA: MIT dissertation.
Haraguchi, Shosuke. (1977) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.
Hardcastle, William J., John Laver, and Fiona E. Gibbon (eds.) (2010) *Handbook of Phonetic Sciences*. Oxford: Wiley-Blackwell, 2nd edition.
Harford, Carolyn and Katherine Demuth. (1999) Prosody outranks syntax: An optimality approach to subject inverstion in Bantu relatives. *Linguistic Analysis* 29, 47–68.
Hargus, Sharon. (1985) *The Lexical Phonology of Sekani*. UCLA dissertation.
Hargus, Sharon and Keren Rice. (2005) *Athabaskan Prosody*. Amsterdam: Benjamins.
Harms, Phillip. (1985) Epena Pedee (Saija) nasalization, in *From Phonology to Discourse: Studies in Six Colombian Languages*, ed. by R. Brend, 13–18, Dallas, TX: Summer Institute of Linguistics.
Harms, Phillip. (1994) *Epena Pedee Syntax*. Studies in the Languages of Colombia 4, Summer Institute of Linguistics and University of Texas at Arlington.
Harms, Robert. (1964) *Finnish Structural Sketch*. Bloomington, IN: Indiana University.
Harrell, F. E. (2001) *Regression Modeling Strategies*. Berlin: Springer.
Harrell, Richard S. (1957) *The Phonology of Colloquial Egyptian Arabic*. American Council of Learned Societies.
Harrington, Jonathan, Sallyanne Palethorpe, and Catherine Watson. (2000) Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association* 30, 63–78.
Harris, Alice. (2000) Where in the word is the Udi clitic? *Language* 76, 593–616.
Harris, Alice. (2002) On the origins of circumfixes in Kartvelian, in *Philologie, Typologie und Sprach-struktur: Festschrift für Winfried Boeder zum 65. Geburtstag*, ed. by W. Bublitz, M. von Roncador, and H. Vater, 305–322, Frankfurt am Main: Peter Lang.
Harris, Alice. (2008a) Explaining exuberant agreement, in *Linguistic Theory and Grammatical Change: The Rosendal Papers*, ed. by T. Eythorsson, Amsterdam: Benjamins.
Harris, Alice. (2008b) On the explanation of typologically unusual structures, in *Language Universals and Language Change*, ed. by Jeff Good, Oxford: Oxford University Press.
Harris, James W. (1967) *Spanish Phonology*. Cambridge, MA: MIT dissertation.
Harris, James W. (1969) *Spanish Phonology*. Cambridge, MA: MIT Press.
Harris, James W. (1983) *Syllable Structure and Stress in Spanish: A Nonlinear Analysis*. Cambridge, MA: MIT Press.

Harris, John. (1985) *Phonological Variation and Change: Studies in Hiberno-English*. Cambridge: Cambridge University Press.
Harris, John. (1990) Segmental complexity and phonological government. *Phonology* 7, 255–301.
Harris, John. (1994) *English Sound Structure*. Oxford: Blackwell.
Harris, John. (1997) Licensing inheritance: An integrated theory of neutralisation. *Phonology* 14, 315–370.
Harris, John and Edmund Gussman. (1998) Final codas: Why the west was wrong, in *Structure and Interpretation: Studies in Phonology*, ed. by Eugeniusz Cyran, 139–162, Lublin: Folium.
Harris, John and Jonathan Kaye. (1990) A tale of two cities: London glottalizing and New York City tapping. *Linguistic Review* 7, 251–274.
Harris, John and Geoff Lindsey. (1995) The elements of phonological representation, in *Frontiers of Phonology: Atoms, Structures, Derivations*, ed. by Jacques Durand and Francis Katamba, 34–79, London and New York: Longman.
Harris, Martin. (1992) French, in *International Encyclopedia of Linguistics*, ed. by William Bright, volume 2, New York: Oxford University Press.
Harris, Zellig S. (1942) Morpheme alternants in linguistic analysis. *Language* 18, 169–180.
Harris, Zellig S. (1951) *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
Harris, Zellig S. (1955) From phoneme to morpheme. *Language* 31, 190–222.
Harrison, K. David. (1999) Vowel harmony and disharmony in Tuvan and Tofa, in *Proceedings of 2nd Asian G.L.O.W.*
Harrison, K. David. (2000) *Topics in the Phonology and Morphology of Tuvan*. Yale University dissertation.
Harrison, K. David, Mark Dras, and Berk Kapicioglu. (2002) Agent-based modeling of the evolution of vowel harmony, in *Proceedings of the 32th Annual Meeting of the North East Linguistic Society*, ed. by Mako Hirotani, 217–236, Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts.
Harrison, K. David and Abigail Kaun. (2000) Pattern-responsive lexicon optimization, in *Proceedings of the 30th Annual Meeting of the North East Linguistic Society*, ed. by M. Hirotani, A. Coetzee, N. Hall, and J. Kim, Amherst, MA: Graduate Linguistic Student Association.
Harrison, Sheldon and Salich Y. Albert. (1976) *Mokilese-English Dictionary*. Honolulu: University Press of Hawaii.
Hartmann, Jutta, Veronika Hegedus, and Henk van Riemsdijk. (2008) *Sounds of Silence: Empty Elements in Syntax and Phonology*. Oxford: Elsevier.
Harvey, Mark and Brett Baker. (2005) Vowel harmony, directionality, and morpheme structure constraints in Warlpiri. *Lingua* 115, 1457–1474.
Haspelmath, Martin. (2004) Does linguistic explanation presuppose linguistic description? *Studies in Language* 28, 554–579.
Haspelmath, Martin. (2007) Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11, 119–132.
Hattori, Shiro. (1961) Prosodeme, syllable structure, and laryngeal phonemes. *Studies in Descriptive and Applied Linguistics* 1, 1–27.
Haugen, Einar. (1938) Notes on "voiced t" in American English. *Dialect Notes* 6, 627–634.
Haugen, Einar. (1956) The syllable in linguistic description, in *For Roman Jakobson*, ed. by Morris Halle *et al.*, 213–221, The Hague: Mouton.
Hawkins, Sarah. (2003) Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, 373–405.
Hawkins, Sarah and R. Smith. (2001) Polysp: A polysystemic, phonetically-

rich approach to speech understanding. *Rivista di Linguistica* 13, 99–188.
Hay, Jennifer and Katie Drager. (2007) Sociophonetics. *Annual Review of Anthropology* 36, 89–103.
Hay, Jennifer, Janet B. Pierrehumbert, and Mary E. Beckman. (2003) Speech perception, well-formedness, and the statistics of the lexicon, in *Papers in Laboratory Phonology VI*, ed. by John Local, Richard Ogden, and Rosalind Temple, 58–74, Cambridge: Cambridge University Press.
Hay, Jennifer B. and Andrea Sudbury. (2005) How rhoticity became /r/-sandhi. *Language* 81, 799–823.
Hayes, Bruce. (1980) *A Metrical Theory of Stress Rules*. Cambridge, MA: MIT dissertation.
Hayes, Bruce. (1985) Iambic and trochaic rhythm in stress rules, in *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society: Parasession on Poetics, Metrics, and Prosody*, 429–446.
Hayes, Bruce. (1986) Inalterability in CV phonology. *Language* 62, 321–351.
Hayes, Bruce. (1989a) Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
Hayes, Bruce. (1989b) The prosodic hierarchy in meter, in *Rhythm and Meter*, ed. by Paul Kiparsky and Gilbert Youmans, 201–260, Orlando, FL: Academic Press.
Hayes, Bruce. (1994) Weight of CVC can be determined by context, in *Perspectives in Phonology*, ed. by Jennifer Cole and Charles Kisseberth, 61–79, Stanford, CA: CSLI.
Hayes, Bruce. (1995a) *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
Hayes, Bruce. (1995b) On what to teach the undergraduates: Some changing orthodoxies in phonological theory, in *Linguistics in the Morning Calm 3*, ed. by Ik-Hwan Lee, 59–77, Seoul: Hanshin.
Hayes, Bruce. (1999) Phonetically driven phonology: The role of Optimality Theory and inductive grounding, in *Functionalism and Formalism in Linguistics*, ed. by Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly, volume 1: General papers, 243–285, Amsterdam: Benjamins.
Hayes, Bruce. (2001) Query Google. Retrieved April 14, 2011 from http://www.linguistics.ucla.edu/people/hayes/.
Hayes, Bruce. (2004) Phonological acquisition in Optimality Theory: The early stages, in *Fixing Priorities: Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 158–203, Cambridge: Cambridge University Press.
Hayes, Bruce. (2009) *Introductory Phonology*. Oxford: Wiley-Blackwell.
Hayes, Bruce and May Abad. (1989) Reduplication and syllabification in Ilokano. *Lingua* 77, 331–374.
Hayes, Bruce, Robert Kirchner, and Donca Steriade (eds.) (2004) *Phonetically Based Phonology*. Cambridge: Cambridge University Press.
Hayes, Bruce and Aditi Lahiri. (1991) Bengali intonational phonology. *Natural Language and Linguistic Theory* 9, 47–96.
Hayes, Bruce and Zsuzsa Cziráky Londe. (2006) Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23, 59–104.
Hayes, Bruce and Donca Steriade. (2004) Introduction: The phonetic bases of phonological markedness, in *Phonetically Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 1–33, Cambridge: Cambridge University Press.
Hayes, Bruce and Tanya Stivers. (2000) Post-nasal voicing. UCLA.
Hayes, Bruce, Bruce Tesar, and Kie Zuraw. (2003) OTSoft 2.1. http://www.linguistics.ucla.edu/people/hayes/otsoft/.
Hayes, Bruce and Colin Wilson. (2008) A maximum entropy model of

phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379–440.
Hayes, Bruce, Kie Suraw, Péter Siptár, and Zsuzsa Londe. (2009) Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85, 822–863.
Hays, D. G. (1964) Dependency theory: A formalism and some observations. *Language* 40, 511–525.
Hayward, Katrina, Richard Hayward, and Salīm Bakhīt Al-Tabūki. (1988) Vowels in Jibbāli verbs. *Bulletin of the School of Oriental and African Studies* 51, 240–250.
Hayward, Richard J. (2006) The OHO constraint, in *Studies in African Linguistic Typology*, ed. by Erhard Voeltz, 155–169, Amsterdam: Benjamins.
Hazen, Timothy J., Kate Saenko, Chia-Hao La, and James R. Glass. (2004) A segment-based audiovisual speech recognizer: Data collection, development, and initial experiments, in *Proceedings of the 6th International Conference on Multimodal Interfaces*, 235–242, Pennsylvania: State College.
Heath, Jeffrey. (1987) *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany, NY: State Univerity of New York Press.
Heath, Jeffrey. (2008) *A Grammar of Jamsay*. Berlin and New York: Mouton de Gruyter.
Heinz, Jeffrey. (2006a) Learning phonotactic patterns from surface forms, in *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, and Michael Scanlon, Somerville, MA: Cascadilla Press.
Heinz, Jeffrey. (2006b) Learning quantity-insensitive stress patterns via local inference, in *Proceedings of the Association for Computational Linguistics Special Interest Group in Phonology 6*, East Stroudsburg, PA: Association for Computational Linguistics.
Heinz, Jeffrey. (2007) *Inductive Learning of Phonotactic Patterns*. UCLA dissertation.
Heinz, Jeffrey, Greg Kobele, and Jason Riggle. (2005) Exploring the typology of quantity insensitive stress systems without gradient constraints. Presented at the annual LSA meeting, Oakland, CA.
Helimski, E. (1998) Nganasan, in *The Uralic Languages*, ed. by D. Abondoloj, 480–515, London: Routledge.
Hellmuth, Sam. (2007) The relation between prosodic structure and pitch accent distribution: Evidence from Egyptian Arabic. *The Linguistic Review* 24, 291–316.
Henderson, Eugénie. (1967) Vowel length and vowel quality in Khasi. *Bulletin of the School of Oriental and African Studies* 30, 564–588.
Henderson, Eugenir. (1992) Khasi clusters and Greenberg's universals. *Mon-Khmer Studies* 18–19, 61–66.
Henderson, John. (1988) *Topics in Eastern and Central Arrernte Grammar*. University of Western Australia dissertation.
Hendricks, Sean. (1999) *Reduplication without Template Constraints: A Study in Bare-Consonant Reduplication*. University of Arizona dissertation.
Hendricks, Sean. (2001) Bare-consonant reduplication without prosodic templates: Expressive reduplication in Semai. *Journal of East Asian Linguistics* 10, 287–306.
Herbert, R. K. (1986) *Language Universals, Markedness Theory, and Natural Phonetic Processes*. Berlin: Mouton de Gruyter.
Herman, Rebecca. (2000) Phonetic markers of global discourse structure in English. *Journal of Phonetics* 28, 466–493.
Herman, Rebecca, Mary E. Beckman, and Kiyoshi Honda. (1996) Subglottal pressure and final lowering in English, in *Proceedings of the 1996 International Conference on Spoken Language Processing*, 145–148.
Hermans, Ben. (2006) Quantity, in *Encyclopedia of Language and Linguistics*, volume 10, 308–314, 2nd edition.
Hetzer, Armin. (1978) *Lehrbuch der Vereinheitlichten Albanischen Schriftsprache*

mit einem Deutsch-Albanischen Wörterbuch. Hamburg: Helmut Buske Verlag.
Hickok, Gregory and David Poeppel. (2000) Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences* 4, 131–138.
Hinskens, Frans. (1992) *Dialect Leveling in Limburg. Structural and Sociolinguistic Aspects*. Catholic University of Nijmegen dissertation.
Hinskens, Frans. (1996) *Dialect Leveling in Limburg: Structural and Sociolinguistic Aspects*. Tübingen: Niemeyer.
Hinskens, Frans, Roeland van Hout, and Leo Wetzels (eds.) (1997) *Variation, Change and Phonological Theory*. Amsterdam: Benjamins.
Hinton, Leanne, Eugene Buckley, Marvin Karmer, and Michael Meacham. (1991) Preliminary analysis of Chacaltongo Mixtec tone, in *Papers From the American Indian Languages Conference*, ed. by James E. Redden, volume 16 of *Occasional Papers in Linguistics*, 147–155, Carbondale, IL: Southern Illinois University.
Hintz, Diane. (2006) Stress in South Conchucos Quechua: A phonetic and phonological study. *International Journal of American Linguistics* 72, 477–521.
Hintzman, Douglas L. (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93, 411–428.
Hirai, Toshio, Norio Higuchi, and Yoshinori Sagisaka. (1997) Comparison of F0 control rules derived from multiple speech databases, in *Computing Prosody: Computational Models for Processing Spontaneous Speech*, ed. by Yoshinori Sagisaka, Nick Campbell, and Norio Higuchi, 211–223, New York: Springer-Verlag.
Hirschberg, Julia and Gregory Ward. (1992) The influence of pitch range, duration, amplitude, and spectral features on the interpretation of the rise-fall-rise intonation contour in English. *Journal of Phonetics* 20, 241–251.
Hjelmslev, Louis. (1939) The syllable as a structural unit, in *Proceedings of the 3rd International Congress of Phonetic Sciences*, 266–272.
Hoard, J. E. (1971) Aspiration, tenseness, and syllabication in English. *Language* 47, 133–140.
Hoard, J. E. (1978) Syllabification in Northwest Indian Languages, with remarks on the nature of syllabic stops and affricates, in *Syllables and Segments*, ed. by Alan Bell and Joan Hooper, 59–72, Amsterdam: North Holland.
Hoberman, Robert. (1988) Emphasis harmony in a Modern Aramaic dialect. *Language* 64, 1–26.
Hochberg, Judith G. (1988a) First steps in the acquisition of Spanish stress. *Journal of Child Language* 15, 273–292.
Hochberg, Judith G. (1988b) Learning Spanish stress: developmental and theoretical perspectives. *Language* 64, 683–706.
Hock, Hans. (1986a) Compensatory lengthening: In defense of the concept "mora." *Folia Linguistica* 20, 431–460.
Hock, Hans. (1986b) *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
Hock, Hans. (1991) *Principles of Historical Linguistics*. Mouton de Gruyter, 2nd edition.
Hockett, Charles F. (1942) A system of descriptive phonology. *Language* 18, 3–21.
Hockett, Charles F. (1954) Two models of grammatical description. *Word* 10, 210–231.
Hockett, Charles F. (1955) *A Manual of Phonology*. Baltimore: Waverly Press, Inc.
Hockett, Charles F. (1958) *A Course in Modern Linguistics*. New York: Macmillan.
Hockett, Charles F. (1960) The origin of speech. *Scientific American* 203, 88–96.
Hoehle, Tilman. (1999) An architecture for phonology, in *Slavic in Head-Driven Phrase Structure Grammar*, ed. by R. Borsley and A. Przepiorkowski, 61–90, Stanford, CL: CSLI.

Hollenbach, Barbara. (1977) Phonetic vs. phonemic correspondence in two Trique dialects, in *Studies in Otomanguean Phonology*, ed. by W. Merrifield, 35–68, Summer Institute of Linguistics.

Hollenbach, Elena Erickson de. (2004) Los tonos del mixteco de Magdalena Penasco. Manuscript, Summer Institute of Linguistics.

Hollich, George, Rochelle Newman, and Peter Jusczyk. (2005) Infants' use of synchronized visual information to separate streams of speech. *Child Development* 76, 598–613.

Holliday, Robyn and Brendan Weekes. (2006) Dissociated developmental trajectories for semantic and phonological false memories. *Memory* 14, 624–636.

Homa, Donald, Deborah Rhoads, and Daniel Chambliss. (1979) Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory* 5, 11–23.

Hombert, Jean-Marie. (1973) Speaking backwards in Bakwiri. *Studies in African Linguistics* 4, 227–236.

Hombert, Jean-Marie (1978) Consonant types, vowel quality, and tone, in *Tone: A Linguistic Survey*, ed. by Viki Fromkin, 77–111, New York: Academic Press.

Hombert, Jean-Marie. (1986) Word games: Some implications for analysis of tone and other phonological constructs, in *Experimental Phonology*, ed. by John J. Ohala and Jeri J. Jaeger, 175–186, San Diego: Academic Press.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. (1979) Phonetic explanations for the development of tones. *Language* 55, 37–58.

Homer, Molly. (1998) *The Role of Contrast in Nasal Harmony*. University of Illinois, Urbana-Champaign dissertation.

Honeybone, Patrick and Ricardo Bermúdez-Otero. (2006) Linguistic knowledge: Perspectives from phonology and from syntax. *Lingua* 116, 543–756.

Hooper, Joan B. (1972) The syllable in phonological theory. *Language* 43, 525–540.

Hooper, Joan B. (1976) *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Hopcroft, John E., Rajeev Motwani, and Jeffrey D. Ullman. (2001) *Introduction to Automata Theory, Languages, and Computation*. Chicago: Addison Wesley Longman.

Hostetler, Roman and Carolyn Hostetler. (1975) A tentative description of Tinputz phonology, in *Phonologies of Five Austronesian Languages*, ed. by Richard Loving, 129–161, Ukarumpa: Summer Institute of Linguistics.

Houlihan, Kathleen and Greg Iverson. (1977) Functionally constrained phonology, in *Current Approaches to Phonological Theory*, ed. by Daniel A. Dinnsen, 50–73, Bloomington, IN: Indiana University Press.

House, Arthur. (1961) On vowel duration in English. *Journal of the Acoustical Society of America* 33, 1174–1178.

House, David. (1990) *Tonal Perception in Speech*. Lund: Lund University Press.

Houston, Derek M., Peter W. Jusczyk, Cecile Kuijpers, Riet Coolen, and Anne Cutler. (2000) Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review* 7, 504–509.

Howard, Irwin. (1972) *A Directional Theory of Rule Application in Phonology*. MIT dissertation.

Hsieh, L., L. B. Leonard, and L. Swanson. (1999) Some differences between English plural noun inflections and third singular verb inflections in the input: The contributions of frequency, sentence position, and duration. *Journal of Child Language* 26, 531–543.

Hualde, José. (1989a) Autosegmental and metrical spreading in the vowel-harmony systems of Northwestern Spain. *Linguistics* 27, 773–805.

Hualde, José. (1989b) The strict cycle condition and noncyclic rules. *Linguistic Inquiry* 20, 675–680.
Hualde, José. (1991) *Basque Phonology*. London: Routledge.
Huddleston, Rodney and Geoffrey K. Pullum. (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Hudson, Grover. (1974) The representation of non-productive alternations, in *Proceedings of the First International Conference on Historical Linguistics*, ed. by John M. Anderson and C. Jones, volume 2, 203–229, Amsterdam: North Holland.
Hudson, Grover. (1975) *Suppletion in the Representation of Alternations*. UCLA dissertation.
Hulst, Harry van der. (1984) *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.
Hulst, Harry van der. (1985) Ambisyllabicity in Dutch, in *Linguistics in the Netherlands 1985*, ed. by H. Bennis and F. Beukema, 57–67, Dordrecht: Foris.
Hulst, Harry van der. (1988a) The dual interpretation of |i|, |u| and |a|, in *Proceedings of the 18th Annual Meeting of the North East Linguistic Society*, ed. by J. Blevins and J. Carter, volume 18, 208–222.
Hulst, Harry van der. (1988b) The geometry of vocalic features, in *Features, Segmental Structure and Harmony Processes (Part II)*, ed. by Harry van der Hulst and Norval Smith, 77–125, Dordrecht: Foris.
Hulst, Harry van der. (1989) Atoms of segmental structure: Components, gestures, and dependency. *Phonology* 6, 253–284.
Hulst, Harry van der. (1995a) The composition of handshapes, in *University of Trondheim, Working Papers in Linguistics*, 1–18, Dragvolll, Norway: University of Trondheim.
Hulst, Harry van der. (1995b) Radical CV phonology: The categorial gesture, in *Frontiers of Phonology: Atoms, Structures, Derivations*, ed. by Jacques Durand and Francis Katamba, 80–116, London and New York: Longman.
Hulst, Harry van der (ed.) (1999) *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter.
Hulst, Harry van der. (2000) Modularity and modality in phonology, in *Phonological Knowledge: Its Nature and Status*, ed. by N. Burton-Roberts, P. Carr, and G. Docherty, 207–244, Oxford: Oxford University Press.
Hulst, Harry van der. (2003) Dutch syllable structure meets Government Phonology, in *A New Century of Phonology and Phonological Theory: A Festschrift for Professor Shosuke Haraguchi on the Occasion of his 60th Birthday*, ed. by Takeru Honma, Masao Okazaki, Toshiyuki Tabata, and Shin-ichi Tanaka, 313–343, Tokyo: Kaitakusha.
Hulst, Harry van der. (2005a) The molecular structure of phonological segments, in *Headhood, Elements, Specification and Contrastivity*, ed. by Philip Carr, Jacques Durand, and Colin Ewen, 193–234, Amsterdam: Benjamins.
Hulst, Harry van der. (2005b) Why phonology is the same, in *The Organization of Grammar: Studies in Honor of Henk van Riemsdijk*, ed. by H. Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 252–262, Berlin and New York: Mouton de Gruyter.
Hulst, Harry van der. (2006a) Dependency phonology, in *The Encyclopedia of Language and Linguistics*, ed. by Keith Brown, volume 3, 451–458, Oxford: Elsevier.
Hulst, Harry van der. (2006b) Licensing in phonology. *The Linguistic Review* 23, 383–428.
Hulst, Harry van der. (2006c) On the parallel organization of linguistic components. *Lingua* 116, 657–688.
Hulst, Harry van der. (2010) A note on recursion in phonology, in *Recursion and*

Human Language, ed. by Harry van der Hulst, 301–342, Berlin: Mouton de Gruyter.
Hulst, Harry van der. (2011) Constraint-based phonologies, in Mirror Lake Papers. University of Connecticut Working Papers 15, ed. by Nilufer Sener, Carlos Buesa Garcia, and Tsuyoshi Swada. Distributed by MIT Working Papers in Linguistics.
Hulst, Harry van der and Nancy A. Ritter. (1998) Kammu minor syllables in head-driven phonology, in *Structure and Interpretation: Studies in Phonology*, ed. by Eugeniusz Cyran, 163–182, Lublin: Folium.
Hulst, Harry van der and Nancy A. Ritter. (1999a) Head-driven phonology, in *The Syllable: Views and Facts*, ed. by Harry van der Hulst and Nancy A. Ritter, 113–167, Berlin, Mouton de Gruyter.
Hulst, Harry van der and Nancy A. Ritter. (1999b) Theories of the syllable, in *The Syllable: View and Facts*, ed. by Harry van der Hulst and Nancy A. Ritter, 13–52, Mouton de Gruyter.
Hulst, Harry van der and Nancy A. Ritter. (2002) Levels, constraints, and heads, in *Asymmetry in Grammar*, ed. by A. M. di Sciullo, volume 2, 147–188, Amsterdam: Benjamins.
Hulst, Harry van der and Norval Smith. (1982) Prosodic domains and opaque segments in autosegmental theory, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 2, 311–336, Dordrecht: Foris.
Hulst, Harry van der and Norval Smith. (1985) Vowel features and umlaut in Djingili, Nyangumarda and Warlpiri. *Phonology Yearbook* 2, 277–303.
Hulst, Harry van der and Norval Smith. (1987) Vowel Harmony in Khalkha and Buriat (East Mongolian), in *Linguistics in the Netherlands 1987*, ed. by Frits Beukema and Peter Coopmans, 81–90, Dordrecht: Foris.
Hulst, Harry van der and Norval Smith (eds.) (1988) *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris.
Hulst, Harry van der and Jeroen van de Weijer. (1995) Vowel harmony, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 495–534, Oxford: Blackwell.
Hulstaert, Gustav. (1961) *Grammaire du lomongo*. Terveuren: Musée Royal de l'Afrique centrale.
Humbert, Helga. (1995) *Phonological Segments: Their Structure and Behaviour*. The Hague: Holland Academic Graphics.
Hume, Elizabeth V. (2003) Language specific markedness: The case of place of articulation. *Studies in Phonetics, Phonology, and Morphology* 9, 295–310.
Hume, Elizabeth V. (2007) Markedness diagnostics and expectation. Presented at the International Conference on Morphology and Phonology, Hanyang University, Seoul, June 2007.
Hume, Elizabeth V. and Keith Johnson. (2001a) A model of the interplay of speech perception and phonology, in *The Role of Speech Perception in Phonology*, ed. by Elizabeth V. Hume and Keith Johnson, 3–26, San Diego, CA: Academic Press.
Hume, Elizabeth V. and Keith Johnson (eds.) (2001b) *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press.
Hume, Elizabeth V., Jennifer Muller, and Aone van Engelenhoven. (1997) Non-moraic geminates in Leti. *Phonology* 14, 371–402.
Hura, Susan, Björn Lindblom, and Randy Diehl. (1992) On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35, 59–72.
Hurlbut, Hope M. (1988) Verb morphology in Eastern Kadazan, in *Pacific Linguistics Series B*, number 97, Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Hyde, Brett. (2002) A restrictive theory of metrical stress. *Phonology* 19, 313–359.
Hyde, Brett. (2007a) Issues in Banawá prosody: Onset sensitivity, minimal words, and syllable integrity. *Linguistic Inquiry* 38, 239–285.
Hyde, Brett. (2007b) Non-finality and weight-sensitivity. *Phonology* 24, 287–334.
Hyman, Larry M. (1970) The role of borrowings in the justification of phonological grammars. *Studies in African Linguistics* 1, 1–48.
Hyman, Larry M. (1975) *Phonology: Theory and Analysis*. New York: Holt, Rinehart, and Winston.
Hyman, Larry M. (1976) Phonologization, in *Linguistic Studies Offered to Joseph Greenberg*, ed. by Alphonse Julliand, volume 2, 407–418, Saratoga, CA: Anna Libri.
Hyman, Larry M. (1977) On the nature of linguistic stress, in *Studies in Stress and Accent*, ed. by Larry M. Hyman, 37–82, Los Angeles: University of Southern California Linguistics Department.
Hyman, Larry M. (1978) Tone and/or accent, in *Elements of Tone, Stress, and Intonation*, ed. by Donna Jo Napoli, 1–20, Washington, DC: Georgetown University Press.
Hyman, Larry M. (1979a) Phonology and noun structure, in *Aghem Grammatical Structure*, ed. by Larry M. Hyman, volume 6 of *Southern California Occasional Papers in Linguistics*, 1–72, Los Angeles: University of Southern California.
Hyman, Larry M. (1979b) A reanalysis of tonal downstep. *Journal of African Languages and Linguistics* 1, 9–29.
Hyman, Larry M. (1981) Tonal accent in Somali. *Studies in African Linguistics* 12, 169–203.
Hyman, Larry M. (1983) Are there syllables in Gokana?, in *Current Approaches to African Linguistics, Vol. 2*, ed. by Jonathan Kaye, Hilda Koopman, Dominique Sportiche, and Andre Dugas, 171–179, Dordrecht: Foris.
Hyman, Larry M. (1985) *A Theory of Phonological Weight*. Dordrecht: Foris.
Hyman, Larry M. (1986) The representation of multiple tone heights, in *The Phonological Representation of Suprasegmentals: Studies on African Languages Offered to John M. Stewart on his 60th Birthday*, ed. by K. Bogers, H. van der Hulst, and M. Mous, 109–152, Dordrecht: Foris.
Hyman, Larry M. (1987) Prosodic domains in Kukuya. *Natural Language and Linguistic Theory* 5, 311–333.
Hyman, Larry M. (1988) Underspecification and vowel height transfer in Esimbi. *Phonology* 5, 255–273.
Hyman, Larry M. (1992a) Moraic mismatches in Bantu. *Phonology* 9, 255–265.
Hyman, Larry M. (1992b) Register tones and tonal geometry, in *The Phonology of Tone: The Representation of Tonal Register*, ed. by Keith Snider and Harry van der Hulst, 75–100, The Hague: Walter de Gruyter.
Hyman, Larry M. (1994) Cyclic phonology and morphology in Cibemba, in *Perspectives in Phonology*, ed. by Jennifer Cole and Charles Kisseberth, 81–112, Stanford, CA: CSLI Publications.
Hyman, Larry M. (1995) Nasal consonant harmony at a distance: The case of Yaka. *Studies in African Linguistics* 24, 5–30.
Hyman, Larry M. (1998) Positional prominence and the "prosodic trough" in Yaka. *Phonology* 15, 41–75.
Hyman, Larry M. (1999) The historical interpretation of vowel harmony in Bantu, in *Bantu Historical Linguistics*, ed. by Jean-Marie Hombert and Larry M. Hyman, Stanford, CA: CSLI Publications.
Hyman, Larry M. (2001a) On the limits of phonetic determinism in phonology: *NC revisited, in *The Role of Speech Perception in Phonology*, ed. by Elizabeth V. Hume and Keith Johnson, 141–185, San Diego, CA: Academic Press.

Hyman, Larry M. (2001b) Privative tone in Bantu, in *Crosslinguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*, ed. by Shigeki Kaji, 237–257, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Hyman, Larry M. (2001c) Tone systems, in *Language Typology and Language Universals: An International Handbook*, ed. by Martin Haspelmath, volume 2, 1367–1380, Berlin and New York: Walter de Gruyter.

Hyman, Larry M. (2002) Is there a right-to-left bias in vowel harmony?, in *Phonologica 2002*, ed. by John R. Rennison, Friedrich Neubarth, and Markus A. Pochtrager, Berlin: Mouton.

Hyman, Larry M. (2003) Sound change, misanalysis, and analogy in the Bantu causative. *Journal of African Languages and Linguistics* 24, 55–90.

Hyman, Larry M. (2005) Elicitation as experimental phonology: Thlangtlang Lai tonology, in *Experimental Approaches to Phonology in Honor of John J. Ohala*, ed. by Maria-Josep Solé, Pam Beddor, and Manjari Ohala, 7–24, Oxford: Oxford University Press.

Hyman, Larry M. (2006a) Affixation by place of articulation: Rare and mysterious, in *University of California Berkeley Phonology Laboratory Annual Report*, University of California Berkeley Linguistics Society.

Hyman, Larry M. (2006b) Word-prosodic typology. *Phonology* 23, 225–257.

Hyman, Larry M. (2007) Universals of tone rules: 30 years later, in *Tones and Tunes I: Studies in Word and Sentence Prosody*, ed. by Tomas Riad and Carlos Gussenhoven, 1–34, Berlin: Mouton De Gruyter.

Hyman, Larry M. (2009) How (not) to do phonological typology: The case of pitch-accent. *Language Sciences* 31, 213–238.

Hyman, Larry M. (2010) Kuki-Thaadow: An African tone system in Southeast Asia, in *Essais de typologie et de linguistique générale*, ed. by Franck Floricic, 31–51, Lyon: Les Presses de l'École Normale Supérieure.

Hyman, Larry M. (2011) Do tones have features?, in *Tones and Features*, ed. by John A. Goldsmith, Elizabeth V. Hume, and Leo Wetzels, The Hague: Mouton.

Hyman, Larry M. and E. Byarushengo. (1984) A model of Haya tonology, in *Autosegmental Studies in Bantu Tone.*, ed. by G. N. Clements and John A. Goldsmith, 53–103, Dordrecht: Foris.

Hyman, Larry M. and Sharon Inkelas. (1997) Emergent templates: The unusual case of Tiene, in *University of Maryland Working Papers in Linguistics: Selected Phonology Papers from the Johns Hopkins Optimality Theory Workshop/Maryland Mayfest*, ed. by V. Miglio and B. Moren, Baltimore: Department of Linguistics, University of Maryland.

Hyman, Larry M., Francis X. Katamba, and Livingstone Walusimbi. (1987) Luganda and the strict layer hypothesis. *Phonology Yearbook* 4, 87–108.

Hyman, Larry M. and William R. Leben. (2000) Suprasegmental processes, in *A Handbook on Inflection and Word Formation*, 587–594, Berlin: Mouton De Gruyter.

Hyman, Larry M. and Armindo Ngunga. (1997) Two kinds of moraic nasals in Ciyao. *Studies in African Linguistics* 26, 165–192.

Hyman, Larry M. and Russell G. Schuh. (1974) Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5, 81–115.

Hyman, Larry M. and Maurice Tadadjeu. (1976) Floating tones in Mbam-Nkam, in *Studies in Bantu tonology*, ed. by Larry M. Hyman, 57–111, Los Angeles: University of Southern California.

Hyman, Larry M. and Kenneth Van Bik. (2004) Directional rule application and

output problems in Hakha Lai tone. *Language and Linguistics* 5, 821–861.
Ichimura, Larry. (2006) *Anti-homophony Blocking and Its Productivity in Transparadigmatic Relations*. Boston University dissertation.
Idsardi, William J. (2000) Clarifying opacity. *The Linguistic Review* 17, 337–350.
Idsardi, William J. (2004) Stress Machine Battlebots. http://www.ling.umd.edu?~idsardi/papers/2004uconnshow.pdf
Idsardi, William J. (2009) Calculating metrical structure, in *Contemporary Views on Architecture and Representations in Phonological Theory*, ed. by Charles Cairns and Eric Raimy, 191–211, Cambridge, MA: MIT Press.
Idsardi, William J. and Eric Raimy. (2005) Remarks on language play. Manuscript, University of Maryland.
Ingleby, Michael and Wiebke Brockhaus. (2002) Phonological primes: Cues and acoustic signatures, in *Phonetics, Phonology, and Cognition*, ed. by Jacques Durand and Bernard Laks, 131–150, Oxford: Oxford University Press.
Ingram, D. (1974) Phonological rules in young children. *Journal of Child Language* 1, 46–64.
Inkelas, Sharon. (1988) Prosodic constrains on syntax: Hausa *Fa*, in *Proceedings of the 7th West Coast Conference on Formal Linguistics*, ed. by Hagit Borer, Palo Alto, CA: Stanford Linguistics Association.
Inkelas, Sharon. (1989) *Prosodic Constituency in the Lexicon*. Stanford University dissertation. Published by Garland Publishing, New York, 1990.
Inkelas, Sharon. (1993) Deriving cyclicity, in *Studies in Lexical Phonology*, ed. by Sharon Hargus and Ellen Kaisse, 75–110, San Diego, CA: Academic Press.
Inkelas, Sharon. (1995) The consequences of optimization for underspecification, in *Proceedings of the 25th Annual Meeting of the North East Linguistic Society*, ed. by E. Buckley and S. Iatridou, 287–302, Amherst, MA: GLSA.
Inkelas, Sharon. (1998) The theoretical status of morphologically conditioned phonology: A case study from dominance, in *Yearbook of Morphology 1997*, 121–155, Berlin: Springer.
Inkelas, Sharon. (1999) Exceptional stress-attracting suffixes in Turkish: Representations vs. the grammar, in *The Prosody-Morphology Interface*, ed. by René Kager, Harry van der Hulst, and Wim Zonneveld, 134–187, Cambridge: Cambridge University Press.
Inkelas, Sharon. (2000) Phonotactic blocking through structural immunity, in *Lexicon in Focus*, ed. by Barbara Stiebels and Dieter Wunderlich, Berlin: Akademie Verlag.
Inkelas, Sharon. (2003) J's rhymes: A longitudinal case study of language play. *Journal of Child Language* 30, 557–581.
Inkelas, Sharon. (2008a) The dual theory of reduplication. *Linguistics* 46, 351–401.
Inkelas, Sharon. (2008b) The morphology-phonology connection, in *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society*, Berkeley, CA: Berkeley Linguistics Society.
Inkelas, Sharon and Cemil Orhan Orgun. (1995) Level ordering and economy in the lexical phonology of Turkish. *Language* 71, 763–793.
Inkelas, Sharon and Cemil Orhan Orgun. (1998) Level (non)ordering in recursive morphology: Evidence from Turkish, in *Morphology and its Relation to Phonology and Syntax*, ed. by Steve LaPointe, Diane Brentari, and Patrick Farrell, 360–392, Stanford, CA: CSLI Publications.
Inkelas, Sharon and Cemil Orhan Orgun. (2003) Turkish stress: A review. *Phonology* 20, 139–161.
Inkelas, Sharon and Draga Zec. (1988) Serbo-Croatian pitch accent. *Languages* 64, 227–248.
Inkelas, Sharon and Draga Zec. (eds.) (1990a) *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.
Inkelas, Sharon and Draga Zec. (1990b) Prosodically constrained syntax, in

The Phonology-Syntax Connection, ed. by Sharon Inkelas and Draga Zec, Chicago: University of Chicago Press.
Inkelas, Sharon and Draga Zec. (1995) The phonology/syntax interface, in *Handbook of Phonological Theory*, ed. by John A. Goldsmith, Oxford: Blackwell.
Inkelas, Sharon and Cheryl Zoll. (2005) *Reduplication: Doubling in Morphology*. Cambridge: Cambridge University Press.
Inkelas, Sharon and Cheryl Zoll. (2007) Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45, 133–171.
IPA. (1949) *The Principles of the International Phonetic Association*. London: University College.
IPA. (1999) *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
Ishihara, Shin. (2007) Major phrase, focus intonation, multiple spell-out. *The Linguistic Review* 24, 137–167.
Ishihara, Shin. (2008) Japanese downstep revisited. Unpublished Manuscript.
Isola, Akinwumi. (1982) Ena: Code-talking in Yoruba. *Journal of West African Languages* 12, 43–51.
Itkonen, Terho. (1957) Mellakoihin vai mellakkoihin? Yleiskielemme eraiden taivutushor-juvuuksien taustaa [Mellakoihin or mellakkoihin? Some background on variability in declension in the standard language]. *Virittäjä* 61, 259–286.
Itô, Junko. (1986) *Syllable Theory in Prosodic Phonology*. University of Massachusetts, Amherst dissertation.
Itô, Junko. (1989) A prosodic theory of epenthesis. *Natural Language and Linguistic Theory 7*, 217–259.
Itô, Junko. (1990) Prosodic minimality in Japanese, in *Papers from the 34th Annual Meeting of the Chicago Linguistic Society*, ed. by Michael Ziolkowski, Manuela Noske, and Karen Deaton, volume 2: *The Parasession on the Syllable in Phonetics and Phonology*, 213–239, Chicago: Chicago Linguistic Society.
Itô, Junko and Jorge Hankamer. (1989) Notes on monosyllabism in Turkish, in *Phonology at Santa Cruz*, ed. by J. Itô and J. Runner, 61–69, Santa Cruz, CA: University of California, Santa Cruz Syntax Research Center.
Itô, Junko, Yoshihisa Kitagawa, and Armin Mester. (1996) Prosodic faithfulness and correspondence: Evidence from a Japanese argot. *Journal of East Asian Linguistics* 5, 217–294.
Itô, Junko and Armin Mester. (1986) The phonology of voicing in Japanese: Theoretical consequences of morphological accessibility. *Linguistic Inquiry* 17, 49–73.
Itô, Junko and Armin Mester. (1995) Japanese phonology: Constraint domains and structure preservation, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 817–838, Oxford: Blackwell.
Itô, Junko and Armin Mester. (1996a) Stem and word in Sino-Japanese, in *Phonological Structure and Language Processing: Cross-Linguistic Studies*, ed. by T. Otake and A. Cutler, 13–44, Berlin: Mouton de Gruyter.
Itô, Junko and Armin Mester. (1996b) Structural economy and OCP interactions in local domains. Presented at the Western Conference on Linguistics, University of California, Santa Cruz.
Itô, Junko and Armin Mester. (1999) Realignment, in *The Prosody-Morphology Interface*, ed. by Kager René, Harry van der Hulst, and Wim Zonnefeld, Cambridge: Cambridge University Press.
Itô, Junko and Armin Mester. (2001a) Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints, in *Proceedings of the 2001 International Conference on Phonology and Morphology*, 3–33, Seoul.

Itô, Junko and Armin Mester. (2001b) The phonological lexicon, in *The Handbook of Japanese Linguistics*, ed. by N. Tsujimura, 62–100, Oxford: Blackwell.

Itô, Junko and Armin Mester. (2002) Lexical and postlexical phonology in Optimality Theory: Evidence from Japanese, in *Linguistische Berichte 11: Resolving Conflicts in Grammars*, ed. by G. Fanselow and C. Féry, 183–207, Hamburg: Buske.

Itô, Junko and Armin Mester. (2003a) *Japanese Morphophonemics: Markedness and Word Structure*. Cambridge, MA: MIT Press.

Itô, Junko and Armin Mester. (2003b) On the sources of opacity in OT: Code processes in German, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, 271–303, Cambridge: Cambridge University Press.

Itô, Junko and Armin Mester. (2003c) Weak layering and word binarity, in *A New Century of Phonology and Phonological Theory: A Festschrift for Professor Shosuke Haraguchi on the Occasion of his 60th Birthday*, ed. by Takeru Honma, Masao Okasaki, Toshiyuki Tabata, and Shin-ichi Tanaka, 26–65, Tokyo: Kaitakusha.

Itô, Junko and Armin Mester. (2006) Prosodic adjunction in Japanese compounds, in *Formal Approaches to Japanese Linguistics 4*, volume 55, 97–112, MIT Working Papers in Linguistics.

Itô, Junko and Armin Mester. (2007) Categories and projections in prosodic structures, in OCP4 (Old World Conference in Phonology), Rhodes, Greece.

Itô, Junko and Armin Mester. (2010) Recursive prosodic phrasing in Japanese, in *Proceedings of the 18th Japanese/Korean Conference*, 147–164, Stanford, CA: CSLI.

Itô, Junko and Armin Mester. (2011) Recursive prosodic phrasing in Japanese, in *Prosody Matters: Essays in Honor of Elisabeth Selkirk*, ed. by Toni Borowsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara, London: Equinox Publishers.

Itô, Junko, Armin Mester, and Jaye Padgett. (1995) Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26, 571–613.

Jackendoff, Ray. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray. (1975) Morphological and semantic regularities in the lexicon. *Language* 51, 639–671.

Jackendoff, R. (1977) *X-bar Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monograph. Cambridge. MA: MIT Press.

Jackendoff, Ray. (1996) *Foundations of Language*. Oxford: Oxford University Press.

Jackendoff, Ray. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford: Oxford University Press.

Jacobs, Haike and Carlos Gussenhoven. (2000) Loan phonology: Perception, salience, the lexicon and OT, in *Optimality Theory: Phonology, Syntax and Acquisition*, ed. by J. Dekkers *et al.*, 193–210, Oxford: Oxford University Press.

Jacobsen, William. (1964) *A Grammar of the Washo Language*. University of California, Berkeley dissertation.

Jacquemot, Charlotte, Christophe Pallier, Denis LeBihan, Stanislas Dehaene, and Emmanuel Dupoux. (2003) Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *The Journal of Neuroscience* 23, 9541–9546.

Jacquemot, Charlotte and Sophie K. Scott. (2006) What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences* 10, 480–486.

Jaeger, T. Florian. (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language* 59, 434–446.

Jaeger, Jeri. (1986) Concept formation as a tool for linguistic research, in *Experimental Phonology*, ed. by John J. Ohala and Jeri J. Jaeger, 211–237, Academic Press.

Jäger, Gerhard. (2007) Maximum entropy models and stochastic Optimality Theory, in *A Festschrift for Joan Bresnan*, ed. by Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson, and Annie Zaenen, 467–479, Stanford, CA: CSLI Publications.

Jahr, Ernst Håkon. (2003) A Norwegian adult language game, anti-language or secret code: The Smoi of Mandal, in *Social Dialectology*, ed. by Peter Trudgill, David Britain, and Jenny Cheshire, 275–286, Amsterdam: Benjamins.

Jakobson, Roman. (1941) *Kindersprache, Aphasie, and Allgemeine Lautgesetze*. The Hague: Mouton. Published in English translation by Mouton, The Hague, 1968.

Jakobson, Roman. (1948) Russian conjugation. *Word* 4, 155–167.

Jakobson, Roman, Gunnar Fant, and Morris Halle. (1952) *Preliminaries to Speech Analysis*. Cambridge, MA: MIT Press.

Jakobson, Roman and Morris Halle. (1956) *Fundamentals of Language*. The Hague: Mouton.

Janas, Petr. (1984) *Niedersorbische Grammatik für den Schulgebrauch*. Bautzen: Domowina-Verlag.

Jantunen, Tommi. (2007) Tavu Suomalaisessa Viittomakielessa (The Syllable in Finnish Sign Language; with English abstract). *Puhe ja kieli* 27, 109–126.

Jantunen, Tommi and Ritva Takkinen. (2010) Syllable structure in sign language phonology, in *Sign Languages: A Cambridge Language Survey*, ed. by Diane Brentari, 312–331, Cambridge: Cambridge University Press.

Jarosz, Gaja. (2006) *Rich Lexicons and Restrictive Grammars – Maximum Likelihood Learning In Optimality Theory*. Johns Hopkins University dissertation.

Jarosz, Gaja. (2007) Restrictiveness in phonological grammar and lexicon learning, in *Papers from the 43rd Annual Meeting of the Chicago Linguistic Society*, 125–139, Chicago: Chicago Linguistic Society.

Jassem, Wiktor and Lutoslawa Richter. (1989) Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17, 317–325.

Javkin, H. (1979) Phonetic universals and phonological change. *Report of the Phonology Laboratory 4, University of California, Berkeley.*

Jenison, Scott, D. and Priscilla B. Jenison. (1991) Obokuitai phonology. *Workpapers in Indonesian Languages and Culture* 9, 69–90.

Jenkins, James J. (1966) Meaningfulness and concepts: Concepts and meaningfulness, in *Analyses of Concept Learning*, ed. by Herbert J. Klausmeier and Chester W. Harris, 65–79, New York: Academic Press.

Jensen, John. (1974) A constraint on variables in phonology. *Language* 50, 675–686.

Jensen, John. (2000) Against ambisyllabicity. *Phonology* 17, 187–235.

Jensen, John and Margaret Stong-Jensen. (1973) Ordering and directionality of iterative rules. *Papers in Linguistics* 6, 66–90.

Jensen, John and Margaret Stong-Jensen. (1979) The relevancy condition and variables in phonology. *Linguistic Analysis* 5, 125–160.

Jensen, Sean. (1994) Is ? an element?: Towards a non-segmental phonology. *SOAS Working Papers in Linguistics and Phonetics* 4, 71–78.

Jesney, Karen. (2007) The locus of variation in weighted constraint grammars. Poster presented at the Workshop on Variation, Gradience and Frquency in Phonology.

Jespersen, Otto. (1904) *Lehrbuch der Phonetik*. Leipzig: Teubner.

Jespersen, Otto. (1922) *Language: Its Nature, Development and Origin*. London: George Allen and Unwin Ltd.

Jessen, Michael and Catherine Ringen. (2002) Laryngeal features in German. *Phonology* 19, 189–218.

Jessner, Ulrike. (2003) The nature of cross-linguistic interaction in the multilingual system, in *The Multilingual Lexicon*, ed. by Jasone Cenoz, Britta Hufeisen, and Ulrike Jessner, 45–55, Dordrecht: Kluwer Academic Publishers.

Jessner, Ulrike. (2006) *Linguistic Awareness in Multilinguals: English As A Third Language*. Edinburgh: Edinburgh University Press.

Jin, Shunde. (1996) *An Acoustic Study of Sentence Stress in Mandarin Chinese*. Ohio State University dissertation.

Johnson, C. Douglas (1972) *Formal Aspects of Phonological Description*. The Hague: Mouton.

Johnson, Keith. (1997) Speech perception without speaker normalization: An exemplar model, in *Talker Variability in Speech Processing*, ed. by Keith Johnson and John W. Mullennix, 145–165, San Diego: Academic Press.

Johnson, Keith. (2004) Massive reduction in conversational American English, in *Spontaneous Speech: Data and Analysis: Proceedings of the 1st Session of the 10th International Symposium*, 29–54, Tokyo, Japan.

Johnson, Mark. (2002) Optimality-theoretic Lexical Functional Grammar, in *The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues*, ed. by Suzanne Stevenson and Paolo Merlo, 59–73, Amsterdam: Benjamins.

Jones, Daniel. (1918) *An Outline of English Phonetics*. Leipzig: Teubner.

Jongstra, W. (2003) *Variation in Reduction Strategies of Dutch Word-initial Consonant Clusters*. University of Toronto dissertation.

Joos, Martin. (1948) *Acoustic Phonetics*. Baltimore: Waverly Press. *Language*, Monograph 23.

Joos, Martin. (1966) *Readings in Linguistics I: The Development of Descriptive Linguistics in America 1925–1956*. Chicago: University of Chicago Press.

José, Brian and Julie Auger. (2005) Geminates and Picard pronominal clitic allomorphy. *Catalan Journal of Linguistics* 4, 127–154.

Jun, Jongho. (1995) *Perceptual and Articulator Factors in Place Assimilation: An Optimality Theoretic approach*. UCLA dissertation.

Jun, Sun-Ah. (1993) *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure*. Ohio State University dissertation. Published by Garland Publishing, New York, 1996.

Jun, Sun-Ah. (2005) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Jun, Sun-Ah and Mary E. Beckman. (1994) Distribution of devoiced high vowels in Korean, in *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 2, 479–482.

Jun, Sun-Ah and Gorka Elordieta. (1997) Intonational structure of Lekeitio Basque, in *Intonation: Theory, Model and Applications*, *Proceedings of the ESCA Workshop*, ed. by Atonis Botinis, Giorgios Kouroupetroglou, and George Carayiannis, 193–196, Athens.

Jun, Sun-Ah and Cécile Fougeron. (2002) Realizations of accentual phrase in French intonation. *Probus* 14, 147–172.

Jurafsky, Daniel and James H. Martin. (2000) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* Englewood Cliffs, NJ: Prentice-Hall. Second edition, 2009.

Jusczyk, Peter W. (2000) *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.

Jusczyk, Peter W., Anne Cutler, and Nancy J. Redanz. (1993a) Preference for the Predominant Stress Patterns of English Words. *Child Development* 64, 675–687.

Jusczyk, Peter W., Angela D. Friederici, Jeanine M. I. Wessels, Vigdis Y. Svenkerud, and Ann Marie Jusczyk. (1993b) Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language* 32, 402–420.

Jusczyk, Peter W., Elizabeth Hohne, and Angela Bauman. (1999) Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics* 61, 1465–1473.

Jusczyk, Peter W., Paul A. Luce, and Jan Charles-Luce. (1994) Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33, 630–645.

Jusczyk, Peter W., Paul Smolensky, and Theresa Allocco. (2002) How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10, 31–73.

Ka, Omar. (1994) *Wolof Phonology and Morphology*. Lanham, MD: University Press of America.

Kabak, Baris and William J. Idsardi. (2007) Perceptual distortions in the adaptation of English consonant clusters: Syllabic structure of consonantal contact constraints? *Language and Speech* 50, 23–52.

Kabak, Baris and Anthi Revithiadou. (2009) An interface approach to prosodic word recursion, in *Phonological Domains: Universals and Deviations*, ed. by J. Grijzenhout and B. Kabak, The Hague: Mouton de Gruyter.

Kabak, Baris and Irene Vogel. (2001) The phonological word and stress assignment in Turkish. *Phonology* 18, 315–360.

Kager, René. (1989) *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: Foris.

Kager, René. (1996) On affix allomorphy and syllable counting, in *Interfaces in Phonology*, ed. by Ursula Kleinhenz, 155–171, Berlin: Akademie Verlag.

Kager, René. (1999) *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, René. (2001) Rhythmic directionality by positional licensing. ROA 514.

Kager, René. (2007) Feet and metrical stress, in *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 195–227, Cambridge: Cambridge University Press.

Kager, René. (2009) Lexical irregularity and the typology of contrast, in *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, Cambridge, MA: MIT Press.

Kager, René, Joe Pater, and Wim Zonneveld. (2004) *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Kahn, Daniel. (1976) *Syllable-Based, Generalizations in English Phonology*. MIT dissertation.

Kahnemuyipour, Arsalan. (2003) Syntactic categories and Persian stress. *Natural Language and Linguistic Theory* 21, 333–379.

Kahnemuyipour, Arsalan. (2004) *The Syntax of Sentential Stress*. University of Toronto dissertation.

Kahnemuyipour, Arsalan. (2009) *The Syntax of Sentential Stress*. Oxford: Oxford University Press.

Kainada, Evia. (2009) *Effects of Prosodic Boundary Strength on Segmental and Suprasegmental Processes in Modern Greek*. University of Edinburgh dissertation.

Kaisse, Ellen M. (1985) *Connected Speech: The Interaction of Syntax and Phonology*. New York: Academic Press.

Kaisse, Ellen M. (1986) Locating Turkish devoicing, in *Proceedings of the 5th Annual Meeting of the West Coast Conference on Formal Linguistics*, 119–128.

Kaisse, Ellen M. and Patricia A. Shaw. (1985) On the theory of lexical phonology, in *Phonology Yearbook*, volume 2, 1–30, Cambridge: Cambridge University Press.

Kakumasu, James. (1986) Urubu-Kaapor, in *Handbook of Amazonian languages*, ed. by Desmond Derbyshire and Geoffrey Pullum, volume 1, 326–406, New York: Mouton.

Kan, Seda. (2009) *Prosodic Domains and the Syntax-Phonology Interface in Turkish*. Master's thesis, Bogazici University.

Kanerva, Jonni. (1989) *Focus and Phrasing in Chichewa Phonology*. Stanford University dissertation.

Kang, Hyeon-Seok. (1994) Variation in past-marking and the question of the system in Trinidadian English, in *Papers from the 30th Annual Meeting of the Chicago Linguistic Society. Volume 2: The Parasession on Variation in Linguistic Theory*, volume 2, 150–164, Chicago: Chicago Linguistic Society.

Kang, Yoonjung. (2003) Perceptual similarity in loanword adaptation: English post-vocalic word-final stops in Korean. *Phonology* 20, 219–273.

Kaplan, Aaron. (2007) Licensing and noniterative harmony in Lango, in *Proceedings of the 37th Annual Meeting of the North East Linguistic Society*, ed. by Emily Elfner and Martin Walkow.

Kaplan, Aaron. (2008) *Noniterativity is an Emergent Property of Grammar*. University of California, Santa Cruz dissertation.

Kaplan, Ronald and Joan Bresnan. (1982) Lexical-Functional Grammar: A formal system for grammatical representation, in *The Mental Representation of Grammatical Relations*, ed. by Joan Bresnan, Cambridge, MA: MIT Press.

Kaplan, Ronald and Martin Kay. (1981 [1994]) Phonological rules and finite-state transducers. Presented to the winter meeting of the Linguistic Society of America, New York. Later published in revised form as "Regular models of phonological rule systems," in *Computational Linguistics* 20(3), 331–378.

Karttunen, Frances. (1970) *Problems of Finnish Phonology*. Indiana University dissertation.

Karttunen, Lauri. (1986) Compilation of two-level phonological rules. Presented at the Annual Meeting of the Linguistic Society of America, San Francisco.

Karttunen, Lauri. (1993) Finite-state constraints, in *The Last Phonological Rule*, ed. by John A. Goldsmith, 173–194, Chicago: University of Chicago Press.

Karttunen, Lauri. (1998) The proper treatment of optimality in computational phonology, in *Proceedings of FSMNLP '98: International Workshop on Finite-State Methods in Natural Language Processing*, 1–12, Ankara, Turkey: Bilkent University.

Karttunen, Lauri, Kimmo Koskenniemi, and Ronald M. Kaplan. (1987) A compiler for two-level phonological rules, in *Tools for Morphological Analysis*, ed. by Mary Dalrymple, Ronald Kaplan, Lauri Karttunen, K. Koskenniemi, S. Shaio, and M. Wescoat, 1–61, Stanford, CA: CSLI.

Katada, Fusa. (1990) On the representation of moras: Evidence from a language game. *Linguistic Inquiry* 21, 641–646.

Kaun, Abigail. (1995) *The Typology of Rounding Harmony: An Optimality Theoretic Approach.* UCLA dissertation.

Kaun, Abigail. (2004) The phonetic foundations of the rounding harmony typology, in *Phonetically Based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 87–116, Cambridge: Cambridge University Press.

Kaun, Abigail and K. David Harrison. (2001) Patterns, pervasive patterns, and feature specification, in *Distinctive Feature Theory*, ed. by T. Alan Hall, 211–236, Berlin: Mouton de Gruyter.

Kavitskaya, Darya. (2002) *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. London: Routledge.

Kavitskaya, Darya and Peter Staroverov. (2010) When an interaction is both opaque and transparent: The paradox of fed counterfeeding. *Phonology* 27, 255–288.

Kawahara, Shigeto. (2002) Similarity among variants: Output-variant correspondence. International Christian University.
Kawahara, Shigeto. (2006) A faithfulness scale projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82, 536–574.
Kawahara, Shigeto. (2007) Sonorancy and geminacy, in *University of Massachusetts Occasional Papers 32*, 145–186.
Kawahara, Shigeto. (2011) The intonation of parentheticals in Japanese, in *Prosody Matters: Essays in Honor of Elisabeth Selkirk*, ed. by Toni Borowsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara, London: Equinox.
Kawahara, Shigeto and Takahito Shinya. (2008) The intonation of gapping and coordination in Japanese: Evidence for intonational phrase and utterance. *Phonetica* 65, 62–105.
Kawakami, Sin. (1957) Jun akusento ni tsuite [On the tone of connected words in Japanese]. *Kokugo Kenky* 7, 44–60.
Kay, Martin. (1987) Nonconcatenative finite-state morphology, in *Proceedings of the Third Conference of the European Chapter of the Association for Computational Linguistics*, 2–10.
Kaye, Jonathan. (1973) Odawa stress and related phenomena, in *Odawa Language Project: Second Report*, ed. by Glyne Piggott and Jonathan Kaye, 42–50, Toronto: University of Toronto Centre for Linguistic Studies.
Kaye, Jonathan. (1974) Opacity and recoverability in phonology. *Canadian Journal of Linguistics* 19, 134–149.
Kaye, Jonathan. (1975) A functional explanation of rule ordering in phonology, in *Papers from the Parasession on Functionalism*, 244–252, Chicago: Chicago Linguistics Society.
Kaye, Jonathan. (1979) On the alleged correlation of markedness and rule function, in *Current Approaches to Phonological Theory*, ed. by Daniel A. Dinnsen, 272–280, Bloomington, IN: Indiana University Press.
Kaye, Jonathan. (1982) Harmony processes in Vata, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 2, 385–452, Dordrecht: Foris.
Kaye, Jonathan. (1988) The phonologist's dilemna: A game theoretic approach to phonological debate. *GLOW Newsletter* 21, 16–19.
Kaye, Jonathan. (1990a) Government in phonology: The case of Moroccan Arabic. *The Linguistic Review* 6, 131–159.
Kaye, Jonathan. (1990b) "Coda" licensing. *Phonology* 7, 301–330.
Kaye, Jonathan. (1992) Do you believe in magic? The story of s+C sequences. *SOAS Working Papers in Linguistics and Phonetics* 2, 293–313. Also in: Henryk Kardela and Bogdan Szymanek (eds.), *A Festschrift for Edmund Gussmann from his friends and colleagues*. Lublin: The Catholic University of Lublin Press. 1996, 155–176.
Kaye, Jonathan. (1995) Derivations and interfaces, in *Frontiers of Phonology: Atoms, Structures, Derivations*, ed. by Jacques Durand and Francis Katamba, 289–332, London and New York: Longman.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. (1985) The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2, 305–328.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. (1990) Constituent structure and government in phonology. *Phonology* 7, 193–231.
Kaye, Jonathan and Glyne Piggott. (1973) On the cyclic nature of Ojibwa *t*-palatalization. *Linguistic Inquiry* 4, 345–362.
Kayne, Richard. (1994) *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
Kazazis, Kostas. (1969) Possible evidence for (near-)underlying forms in the

speech of a child, in *Papers from the 5th Annual Meeting of the Chicago Linguistic Society*, 382–386, Chicago: Chicago Linguistic Society.
Kaze, Jeffery. (1989) *Metaphony in Spanish and Italian Dialects Revisited*. University of Illinois, Urbana-Champaign dissertation.
Kean, Mary-Louise. (1974) The strict cycle in phonology. *Linguistic Inquiry* 5, 179–203.
Keating, Patricia A. (1985) Universal phonetics and the organization of grammars, in *Phonetic Linguistics*, ed. by Victoria A. Fromkin, 115–132, San Diego, CA: Academic Press.
Keating, Patricia A. (1990) Phonetic representations in a generative grammar. *Journal of Phonetics* 18, 321–334.
Keating, Patricia A. (1998) Word-level phonetic variation in large speech corpora, in *ZAS Papers in Linguistics 11*, ed. by A. Alexiadou, N. Fuhrop, U. Kleinhenz, and P. Law, 35–50, Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung.
Keating, Patricia A., Taehong Cho, Cécile Fougeron, and Chai-Shune Hsu. (2003) Domain-initial strengthening in four languages, in *Papers in Laboratory Phonology VI*, ed. by John Local, Richard Ogden, and Rosalind Temple, 145–163, Cambridge: Cambridge University Press.
Keer, Edward. (1999) *Geminates, the OCP and the nature of CON*. Rutgers University dissertation.
Kegl, Judy. (1985) *Locative Relations in ASL Word Formation, Syntax, and Discourse*. Cambridge, MA: MIT dissertation.
Kegl, Judy, Ann Senghas, and Marie Coppola. (1999) Creation through contact: Sign language emergence and sign language change in Nicaragua, in *Language Creation and Language Change*, ed. by M. DeGraff, 179–238, Cambridge, MA: MIT Press.
Kehoe, M., G. Hilaire-Debove, K. Demuth, and C. Lleó. (2008) The structure of branching onsets and rising diphthongs: Evidence from the acquisition of French and Spanish. *Language Acquisition: A Journal of Developmental Linguistics* 15, 53.
Kehoe, M. and C. Stoel-Gammon. (2001) Development of syllable structure in English-speaking children with particular reference to rhymes. *Journal of Child Language* 28, 393–432.
Kelepir, Meltem. (1999) Turkish emphatic reduplication and antifaithfulness, in *Proceedings of ConSOLE 7*, University of Bergen, Norway.
Kelkar, Ashok R. (1968) *Studies in Hindi-Urdu I: Introduction and Word, Phonology*. Poona: Deccan College.
Keller, Frank. (2000) *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. University of Edinburgh dissertation.
Keller, Frank. (2006) Linear Optimality Theory as a model of gradience in Grammar, in *Gradience in Grammar: Generative Perspectives*, ed. by Gisbert Fanselow, Caroline Féry, Ralph Vogel, and Matthias Schlesewsky, 270–287, Oxford: Oxford University Press.
Keller, Frank and Ash Asudeh. (2002) Probabilistic learning algorithms and Optimality Theory. *Linguistic Inquiry* 33, 225–244.
Keller, Natalia Eraso. (1999) Cambios tonales en la palabra verbal de la lengua tanimuca, in *Congresso de Lingüística Amerindia y Criolla*, 72–95.
Kemler Nelson, Deborah G., Peter W. Jusczyk, Denise R. Mandel, James Myers, Alice Turk, and LouAnn Gerken. (1995) The headturn preference procedure for testing auditory perception. *Infant Behavior and Development* 18, 111–116.
Kennedy, George A. (1953) Two tone patterns in Tangsic. *Language* 29, 367–373.
Kennedy, Robert. (2003) *Confluence in Phonology: Evidence from Micronesian Reduplication*. University of Arizona dissertation.

Kenstowicz, Michael. (1994a) *Phonology in Generative Grammar*. Oxford: Blackwell.
Kenstowicz, Michael. (1994b) Sonority-driven stress. ROA 33.
Kenstowicz, Michael. (1996) Base identity and uniform exponence: Alternatives to cyclicity, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks, 363–394, Salford: ESRI.
Kenstowicz, Michael. (1997) Quality-sensitive stress. *Rivista di Linguistica* 9, 157–188.
Kenstowicz, Michael. (2003) The role of perception in loanword phonology. *Studies in African Linguistics* 32, 95–112.
Kenstowicz, Michael. (2005) Paradigmatic uniformity and contrast, in *Paradigms in Phonological Theory*, ed. by Laura Downing, T. Alan Hall, and R. Raffelsiefen, 145–169, Oxford: Oxford University Press.
Kenstowicz, Michael. (2007) Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences* 29, 316–340.
Kenstowicz, Michael and Charles Kisseberth. (1979) *Generative Phonology*. New York: Academic Press.
Kenstowicz, Michael and Charles W. Kisseberth. (1971) Unmarked bleeding orders, in *Studies in the Linguistic Sciences 1*, University of Illinois Urbana-Champaign.
Kenstowicz, Michael and Charles W. Kisseberth. (1977) *Topics in Phonological Theory*. New York: Academic Press.
Kenstowicz, Michael and Charles W. Kisseberth. (1990) Chizigula tonology: The word and beyond, in *The Phonology-Syntax Connection*, ed. by Sharon Inkelas and Draga Zec, Chicago: University of Chicago Press.
Kenstowicz, Michael and Charles Pyle. (1973) On the phonological integrity of geminate clusters, in *Issues in Phonological Theory*, ed. by Michael Kenstowicz and Charles Kisseberth, 27–43, The Hague: Mouton.
Kent, Raymond D. (1984) Psychobiology of speech development: Coemergence of language and a movement system. *American Journal of Physiology* 246, 888–894.
Kenyon, John S. and Thomas A. Knott. (1953) *A Pronunciation Dictionary of American English*. Springfield, MA: Merriam, 4th edition.
Kerkhoff, Annemarie. (2007) *Acquisition of Morpho-Phonology: The Dutch Voicing Alternation*. University of Utrecht dissertation.
Keuleers, Emmanuel (2008) *Memory-based Learning of Inflectional Morphology*. Antwerp: University of Antwerp.
Keune, Karen, Roland van Hout, and R. H. Baayen. (2005) Social, geographical, and register variation in Dutch: From written "mogelijk" to spoken "mok." *Corpus Linguistics and Linguistic Theory* 1, 183–223.
Key, Harold. (1961) Phonotactics of Cayuvava. *International Journal of American Linguistics* 27, 143–150.
Key, Harold. (1967) *Morphology of Cayuvava*. The Hague: Mouton.
Key, Mary. (1968) *Comparative Tacanan Phonology with Cavineña Phonology and Notes on Pano-Tacanan Relationship*. The Hague: Mouton.
Keynon, J. S. and T. A. Knott. (1953) *A Pronouncing Dictionary of American English*. Springfield, MA: Merriam.
Keyser, Samuel J. and Paul Kiparsky. (1984) Syllable structure in Finnish phonology, in *Language Sound Structure*, ed. by Mark Aronoff and Richard Oehrle, 7–31, Cambridge, MA: MIT Press.
Keyser, Samuel J. and Kenneth N. Stevens. (2006) Enhancement and overlap in the speech chain. *Language* 82, 33–63.
Keyser, Samuel Jay and Kenneth N. Stevens. (2001) Enhancement revisited, in *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 271–291, Cambridge, MA: MIT Press.

Khumalo, James. (1987) *An Autosegmental Account of Zulu Phonology*. University of Witwatersrand dissertation.
Kidda, Mario. (1985) *Tangale Phonology: A Descriptive Analysis*. Berlin: Dietrich Reimer Verlag.
Kim, Eui-Jung. (1996) Edopi phonology. Manuscript SIL.
Kim, Hyunsoon and Allard Jongman. (1996) Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics* 24, 295–312.
Kim, Yuni. (2008) *Topics in the Phonology and Morphology of San Francisco del Mar Huave*. University of California, Berkeley dissertation.
Kim-Renaud, Young-Key. (1986) *Studies in Korean Linguistics*. Seoul: Hanshin Publishing Company.
Kimper, Wendell A. (2011) Locality and globality in phonological variation. *Natural Language and Linguistic Theory* 29.
Kimenyi, Alexandre. (1979) *Studies in Kinyarwanda and Bantu Phonology*. Carbondale, IL: Linguistic Research Inc.
Kindaichi, Haruhiko. (1957) Nihongo akusento no hiken [On the unit of Japanese accent]. *Kokugo Kenky* 7, 1–32.
King, Robert D. (1969) *Historical Linguistics and Generative Grammar*. Englewood Cliffs, NJ: Prentice-Hall.
King, Robert D. (1973) Rule insertion. *Language* 29, 551–578.
King, Robert D. (1980) The history of final devoicing in Yiddish, in *The Field of Yiddish: Studies in Language, Folklore, and Literature, Fourth Collection*, ed. by M. I. Herzog, D. Miron B. Kirshenblatt-Gimblett, and R. Wisse, 371–430, Philadelphia, PA: Institute for the Study of Human Issues.
King, S. and P. Taylor. (2000) Detection of phonological features in continuous speech using neural networks. *Computer Speech and Language* 14, 333–353.
Kingdon, Roger. (1939) Tonetic stress marks for English. *Le maître phonétique* 54, 60–64.
Kingston, John. (2003) Mechanisms of tone reversal, in *Cross-linguistic Studies of Tonal Phenomena*, ed. by Shigeki Kaji, 57–120, Tokyo: ILCAA.
Kingston, John. (2005) The phonetics of Athabaskan tonogenesis, in *Athabaskan Prosody*, ed. by Sharon Hargus and Keren Rice, 137–184, Benjamins.
Kingston, John and Mary E. Beckman (eds.) (1990) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.
Kingston, John and Randy L. Diehl. (1994) Phonetic knowledge. *Language* 70, 419–454.
Kiparsky, Paul. (1965) *Phonological Change*. MIT dissertation.
Kiparsky, Paul. (1968a) How abstract is phonology? Distributed by Indiana University Linguistics Club.
Kiparsky, Paul. (1968b) Linguistic universals and linguistic change, in *Universals in Linguistic Theory*, ed. by Emmon Bach and Robert T. Harms, 170–202, New York: Holt, Reinhart, and Winston.
Kiparsky, Paul. (1971) Historical linguistics, in *A Survey of Linguistic Science*, ed. by W. O. Dingwall, 576–642, College Park, MD: University of Maryland Linguistics Program.
Kiparsky, Paul. (1972) Explanation in phonology, in *Goals in Linguistic Theory*, ed. by S. Peters, 189–227, Englewood Cliffs, NJ: Prentice-Hall.
Kiparsky, Paul. (1973a) "Elsewhere" in phonology, in *A Festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, New York: Holt, Reinhart, and Winston.
Kiparsky, Paul. (1973b) The inflectional accent in Indo-European. *Language* 49, 794–849.
Kiparsky, Paul. (1973c) Phonological representations, in *Three Dimensions of Linguistic Theory*, ed. by Osamu Fujimura, 3–136, Tokyo: TEC.

Kiparsky, Paul. (1976) Abstractness, opacity, and global rules, in *The Application and Ordering of Grammatical Rules*, ed. by Andreas Koutsoudas, 160–184, The Hague: Mouton.

Kiparsky, Paul. (1979) Metrical structure assignment is cyclic. *Linguistic Inquiry* 10(3), 421–441.

Kiparsky, Paul. (1981a) Remarks on the metrical structure of the syllable, in *Phonologica*, ed. by Wolfgang U. Dressler, Oskar E. Pfeiffer, and John E. Rennison, Innsbruck.

Kiparsky, Paul. (1981b) Vowel harmony. Unpublished manuscript.

Kiparsky, Paul. (1982a) *Explanation in Phonology*. Dordrecht: Foris.

Kiparsky, Paul. (1982b) From cyclic phonology to lexical phonology, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 1, 131–175, Dordrecht: Foris.

Kiparsky, Paul. (1982c) Lexical phonology and morphology, in *Linguistics in the Morning Calm*, ed. by I. S. Yang, 3–91, Seoul: Hanshin.

Kiparsky, Paul. (1984) On the lexical phonology of Icelandic, in *Nordic Prosody II*, ed. by C.-C. Elert, I. Johansson, and E. Strangert, 135–162, Umeå: University of Umeå.

Kiparsky, Paul. (1985) Some consequences of lexical phonology. *Phonology Yearbook* 2, 83–136.

Kiparsky, Paul. (1993a) Blocking in non-derived environments, in *Studies in Lexical Phonology*, ed. by Sharon Hargus and Ellen Kaisse, 277–313, San Diego, CA: Academic Press.

Kiparsky, Paul. (1993b) An OT perspective on phonological variation. Presented at Rutgers Optimality Workshop 1993 and at NWAV 1994, Stanford University.

Kiparsky, Paul. (1995) The phonological basis of sound change, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 640–670, Oxford: Blackwell.

Kiparsky, Paul. (2000) Opacity and cyclicity. *The Linguistic Review* 17, 351–367.

Kiparsky, Paul. (2003a) Finnish noun inflection, in *Generative Approaches to Finnic Linguistics*, ed. by Diane Nelson and Satu Manninen, Stanford: CSLI.

Kiparsky, Paul. (2003b) Syllables and moras in Arabic, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, 147–182, Cambridge: Cambridge University Press.

Kiparsky, Paul. (2006) The amphichronic program vs. evolutionary phonology. *Theoretical Linguistics* 32, 217–236.

Kiparsky, Paul. (2008a) Fenno-Swedish quantity: contrast in stratal OT, in *Rules, Constraints, and Phonological Phenomena*, ed. by Bert Vaux and Andrew Nevins, 185–220, Oxford: Oxford University Press.

Kiparsky, Paul. (2008b) Universals constrain change; change results in typological generalizations, in *Language Universals and Language Change*, ed. by Jeff Good, 23–53, Oxford: Oxford University Press.

Kiparsky, Paul. (2008c) Weight and length. Presented at CUNY Conference on the Syllable. New York City, January 17–19.

Kiparsky, Paul. (2011) Compensatory lengthening, in *Handbook of the Syllable*, ed. by C. Cairns and E. Remy, 33–69, Leiden: Brill.

Kiparsky, Paul, (to appear a) *Paradigms and Opacity*. Stanford, CA: CSLI.

Kiparsky, Paul, (to appear b) Reduplication in stratal OT, in *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, ed. by L. Uyechi and L. H. Wee, Stanford, CA: CSLI.

Kiparsky, Paul and Morris Halle. (1977) Towards a reconstruction of the Indo-European accent, in *Studies in Stress and Accent*, ed. by Larry M. Hyman, 209–238, Los Angeles: Department of Linguistics, University of Southern California.

Kiparsky, Paul and Lise Menn. (1977) On the acquisition of phonology, in *Language Learning and Thought*, ed. by J. Macnamara, 47–78, New York: Academic Press.

Kirby, James P. (2010) *Cue Selection and Category Restructuring in Sound Change*. University of Chicago dissertation.

Kirby, James P. and Alan C. L. Yu. (2007) Lexical and phonotactic effects on wordlikeness judgements in Cantonese, in *Proceedings of the International Congress of the Phonetic Sciences XVI*, 1161–1164.

Kirby, James P. (to appear a) Acquisition of covert contrasts: An unsupervised learning approach, in *Proceedings of the 46th Annual Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society.

Kirby, James P. (to appear b) The role of probabilistic enhancement in phonologization, in *Origins of Sound Change: Approaches to Phonologization*, ed. by Alan C. L. Yu, Oxford: Oxford University Press.

Kirchner, Robert. (1993) Turkish vowel harmony and disharmony: An optimality theoretic account. Presented at the Rutgers Optimality Workshop I.

Kirchner, Robert. (1996) Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27, 341–349.

Kirchner, Robert. (1997) Contrastiveness and faithfulness. *Phonology* 14, 83–111.

Kirchner, Robert and Roger Moore. (2008) Computing phonological generalization over real speech exemplars. Manuscript, University of Alberta and University of Sheffield.

Kirk, Cecilia. (2008) Substitution errors in the production of word-initial and word-final consonant clusters. *Journal of Speech, Language and Hearing Research* 51, 1–14.

Kirk, Cecilia and Katherine Demuth. (2005) Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language* 32, 709–734.

Kirk, Cecilia and Katherine Demuth. (2006) Accounting for variability in 2-year-olds' production of coda consonants. *Language Learning and Development* 2, 97–118.

Kisseberth, Charles W. (1969) *Theoretical Implications of Yawelmani Phonology*. University of Illinois Champaign-Urbana dissertation.

Kisseberth, Charles W. (1970) On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291–306.

Kisseberth, Charles W. (1973) On the alternation of vowel length in Klamath: A global rule, in *Issues in Phonological Theory*, ed. by M. Kenstowicz and C. Kisseberth, The Hague: Mouton.

Kisseberth, Charles W. (1976) The interaction of phonological rules and the polarity of language, in *The Application and Ordering of Grammatical Rules*, ed. by Andreas Koutsoudas, 41–54, The Hague: Mouton.

Kisseberth, Charles W. (1984) Digo tonology, in *Autosegmental Studies in Bantu Tone*, ed. by G. N. Clements and John A. Goldsmith, 105–182, Dordrecht: Foris.

Kisseberth, Charles W. (1994) On domains, in *Perspectives in Phonology*, ed. by Jennifer Cole and Charles Kisseberth, 133–166, Stanford, CA: CSLI.

Kisseberth, Charles W. (2005) Accent and phrasing in Chimwiini, in *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone and Descriptive Studies*, ed. by Shigeki Kaji, 129–145, Tokyo, Tokyo University of Foreign Studies. Research ILCAA.

Kisseberth, Charles W. and Mohammad Imam Abasheikh. (1974) Vowel length in Chimwiini: A case study of the role of grammar in phonology, in *Papers from the Parasession on Natural Phonology*, ed. by A. Bruck *et al.*, 193–209, Chicago: Chicago Linguistic Society.

Kisseberth, Charles W. and Erez Volk. (2007) Theoretical consequences of depressor consonants. Presented at the Workshop on Segments and Tone, Meertens Instituut, Amsterdam.

Kitto, Catherine and Paul de Lacy. (1999) Correspondence and epenthetic quality, in *Austronesian Formal Linguistics Association* 6, 181–200, Toronto Working Papers in Linguistics.

Klatt, Dennis. (1979) Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics* 7, 279–312.

Klein, Ewen. (1987) Towards a declarative phonology. Handout, Edinburgh University, November 10, 1987.

Klein, Thomas. (1995) *Umlaut in Optimality Theory*. University of Delaware dissertation.

Klein, Thomas. (2003) Syllable structure and lexical markedness in creole morphophonology: Determiner allomorphy in Haitian and elsewhere, in *The Phonology and Morphology of Creole Languages*, ed. by I. Plag, 209–228, Tübingen: Max Niemeyer.

Klima, Edward and Ursula Bellugi. (1979) *The Signs of Language*. Cambridge, MA: Harvard University Press.

Knuth, D. E. (1973) *The Art of Computer Programming. Vol. 3: Sorting and Searching*. Reading, MA: Addison-Wesley.

Kochetov, Andrei and Marianne Pouplier. (2009) Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. *Phonology* 25, 399–431.

Kohler, Klaus J. (1990) Segmental reduction in connected speech in German: Phonological effects and phonetic explanations, in *Speech Production and Speech Modelling*, ed. by William Hardcastle and Alain Marchal, 69–92, Dordrecht: Kluwer.

Kohlrausch, Armin, Dirk Püschel, and Henning Alphei. (1992) Temporal resolution and modulation analysis in models of the auditory system, in *The Auditory Processing of Speech: From Sounds to Words*, ed. by M. E. H. Schouten, 85–98, Berlin: Mouton de Gruyter.

Kontra, Miklós, Catherine O. Ringen and Joseph P. Stemberger. (1991) The effect of context on suffix vowel choice in Hungarian vowel harmony, in *Proceedings of the 14th International Congress of Linguists*, ed. by Werner Bahner, Joachim Schildt, and Dieter Viehweger, volume 1, 450–453, Berlin: Akademie-Verlag.

Koo, Hahn and Jennifer Cole. (2006) On learnability and naturalness as constraints on phonological grammar, in *Proceedings of the ISCA Workshop on Experimental Linguistics*, 165–168.

Koo, Hahn and Jennifer Cole. (2007) Gradient perceptual facilitation from phonotactic knowledge, in *Papers from the 43th Annual Meeting of the Chicago Linguistic Society*, 1, 167–181, Chicago: Chicago Linguistic Society.

Kooij, Els. (2002) *Reducing Phonological Categories in Sign Language of the Netherlands: Phonetic Implementation and Iconic Motivation*. Utrecht: LOT (Netherlands Graduate School of Linguistics).

Kori, Shiro. (1987) The tonal behavior of Osaka Japanese: An interim report, in *Ohio State University Working Papers in Linguistics*, volume 36, 31–61, Ohio State University.

Kormushin, Igor Valentinovich. (1998) *Udykheiskii (Udegeiskii) Iazyk*. Moscow: Nauka.

Kornai, András. (1994) Relating phonetic and phonological categories, in *Language Computations*, ed. by E. S. Ristad, volume 17, 21–36, Philadelphia, PA: American Mathematical Society.

Koskenniemi, Kimmo. (1983) *Two-level Morphology: A General Computational Model for Word-form Recognition and Production*. Helsinki: University of

Helsinki Department of General Linguistics.
Koskenniemi, Kimmo. (1984) Two-level morphology: A general computational model for wordform recognition and production, in *Proceedings of the 10th International Conference on Computational Linguistics*, 178–181.
Kossmann, Maarten and Harry Stroomer. (1997) Berber phonology, in *Phonologies of Asia and Africa*, ed. by A. Kaye, 461–475: Eisenbrauns.
Koutsoudas, Andreas, Gerald Sanders, and Craig Noll. (1974) The application of phonological rules. *Language* 50, 1–28.
Kowalski, T. (1929) *Karaimische Texte im Dialekt von Troki*. Krakow: Polska Akademia Umiejtności.
Kraehenmann, Astrid. (2001) *Quantity and Prosody Asymmetries in Alemannic: Synchronic and Diachronic perspectives*. University of Konstanz dissertation. Published by Mouton de Gruyter, Berlin, 2003.
Kraft, K. and M. Kraft. (1973) *Introductory Hausa*. Berkeley, CA: University of California Press.
Krämer, Martin. (2001) Yucatec Maya vowel alternation–harmony as syntagmatic identity. *Zeitschrift für Sprachwissenschaft* 20, 175–217.
Krämer, Martin. (2003) *Vowel Harmony and Correspondence Theory*. New York: Mouton de Gruyter.
Kratochvil, Paul. (1998) Intonation in Beijing Chinese, in *Intonation Systems: A Survey of Twenty Languages*, ed. by Daniel Hirst and Albert Di Cristo, 417–430, Cambridge: Cambridge University Press.
Kratzer, Angelika and Elisabeth Selkirk. (2007) Phase theory and prosodic spellout: The case of verbs.
Kratzer, Angelika and Elisabeth Selkirk. (2010) Distinguishing contrastive, given and new information. 24, 93–136. Presented at International Conference on Information Structure, Potsdam, July, 2010.
Krauss, Michael E. (1975) Chilcotin phonology: A descriptive and historical report, with recommendations for a Chilcotin orthography. Manuscript, Alaska Native Language Center.
Krauss, Scott. (1981) *Topics in Chukchee Phonology and Morphology*. University of Illinois Urbana-Champaign dissertation.
Krishnamurti, Bhadriraju. (1998) Regularity of sound change through lexical diffusion: A study of s > h >> in Gondi dialects. *Language Variation and Change* 10, 193–220.
Kroch, Anthony. (1989) Reflexes of grammar in patterns of language changes. *Language Variation and Change* 1, 199–244.
Kroeker, Menno H. (1977) The role of tone in Nambikuára. *Arquivos de Anatomia e Antropologia* 2, 119–143.
Krott, Andrea, R. Harald Baayen, and Robert Schreuder. (2001) Analogy in morphology: Modeling the choice of linking morphemes in Dutch. *Linguistics* 39, 51–93.
Krueger, John Richard. (1961) *Chuvash Manual*. Bloomington, IN: Indiana University Press.
Krueger, John Richard. (1962) *Yakut Manual*. Bloomington, IN: Indiana University Publications in Uralic and Altaic Series.
Kubozono, Haruo. (1988) *The Organisation of Japanese Prosody*. University of Edinburgh dissertation.
Kubozono, Haruo. (1989a) The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech* 32, 249–278.
Kubozono, Haruo. (1989b) Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6, 39–67.
Kubozono, Haruo. (1993) *The Organization of Japanese Prosody*. Tokyo: Kurosio.
Kubozono, Haruo. (1999) Mora and syllable, in *The Handbook of Japanese Linguistics*, ed. by Natsuko Tsujimura, 31–61, Oxford: Blackwell.

Kubozono, Haruo. (2007) Focus and intonation in Japanese, in *Interdisciplinary Studies on Information Structure 9: Proceedings of the 2nd Workshop on Prosody, Syntax, and Information Structure*, ed. by Shin Ishihara, Potsdam, University of Potsdam.

Kuperman, Victor, Rob Schreuder, Raymond Bertram, and R. Harald Baayen. (2008) Reading of multimorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 62, 83–97.

Kupin, Joseph J. (1982) *Tongue Twisters as a Source of Information about Speech Production*. Indiana University Linguistics Club.

Kurisu, Kazutaka. (2001) *The Phonology of Morpheme Realization*. University of California, Santa Cruz dissertation.

Kuroda, S.-Y. (1967) *Yawelmani Phonology*. Cambridge, MA: MIT Press.

Kuryłowicz, Jerzy. (1948) Contribution à la théorie de la syllabe. *Bulletin de la Societé polonaise de linguistique* 8, 80–104.

Kuryłowicz, Jerzy. (1952) Uwagi o polskich grupach spółgłoskowych [Remarks on Polish consonantal groups]. *Biuletyn Polskiego Towarzystwa Językoznawczego* 12, 221–232.

Kučera, Henry. (1961) *The Phonology of Czech*. 's-Gravenhage: Mouton.

Kwenzi Mikala, Jérôme. (1980) Esquisse phonologique du punu, in *Elements de description du punu*, ed. by F. Nsukua-Nkutsi, 7–18, CRLS Univérsité Lyon II.

Labov, William. (1969) Contraction, deletion and inherent variability of the English copula. *Language* 45, 715–762.

Labov, William. (1971) Methodology, in *A Survey of Linguistic Science*, ed. by William Orr Dingwall, 412–497, University of Maryland Linguistics Program.

Labov, William. (1989) The child as linguistic historian. *Language Variation and Change* 1, 85–97.

Labov, William. (1994) *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Labov, William. (1997) Resyllabification, in *Variation, Change, and Phonological Theory*, ed. by Frans Hinskens, Roeland van Hout, and Leo Wetzels, 145–179, Amsterdam: Benjamins.

Labov, William. (2004) Quantitative analysis of linguistic variation, in *Sociolinguistics: An International Handbook of the Science of Language and Society*, ed. by Ulrich Ammon, Norbert Dittmer, Klaus J. Mattheier, and Peter Trudgill, 6–21, Berlin: De Gruyter, 2nd edition.

Labov, William, Sharon Ash, and Charles Boberg. (2006) *Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Mouton de Gruyter.

Labov, William, Mark Karen and Corey Miller. (1991) Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3, 33–74.

Labov, William, Malcah Yaeger, and Richard Steiner. (1972) *A Quantitative Study of Sound Change in Progress*. US Regional Survey.

LaCharité, Darlene and Carole Paradis. (2005) Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36, 223–258.

Ladd, D. Robert. (1980) *The Structure of Intonational Meaning: Evidence from English*. Bloomington, IN: Indiana University Press.

Ladd, D. Robert. (1986) Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3, 311–340.

Ladd, D. Robert. (1988) Declination "reset" and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84, 538–544.

Ladd, D. Robert. (1992) An introduction to intonational phonology, in *Papers in Laboratory Phonology II*, ed. by Gerard J. Docherty and D. Robert Ladd, 321–334, Cambridge: Cambridge University Press.

Ladd, D. Robert. (1993) In defense of a metrical theory of intonational downstep, in *The Phonology of Tone: The Representation of Tonal Register*, ed. by Harry van der Hulst and Keith Snider, 109–132, Dordrecht: Foris.

Ladd, D. Robert. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, D. Robert. (2006) "Distinctive phones" in surface representation, in *Laboratory Phonology VIII*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 1–26, Berlin: Mouton de Gruyter.

Ladd, D. Robert. (2008a) *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, D. Robert. (2008b) Review of *Prosodic Typology*, by Sun-Ah Jun (ed.) *Phonology* 25, 372–376.

Ladd, D. Robert, Ineke Mennen, and Astrid Schepman. (2000) Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107, 2685–2696.

Ladd, D. Robert and Astrid Schepman. (2003) "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics* 31, 81–112.

Ladd, D. Robert and James M. Scobbie. (2003) External sandhi as gestural overlap? Counter-evidence from Sardinian, in *Papers in Laboratory Phonology VI: Phonetic Interpretation*, ed. by John Local, Richard Ogden, and Rosalind Temple, 164–182, Cambridge: Cambridge University Press.

Ladefoged, Peter. (1971) *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter. (1990) Some reflections on the IPA. *Journal of Phonetics* 18, 335–346.

Ladefoged, Peter. (1993) *A Course in Phonetics*. Fort Worth, TX: Harcourt Brace Jovanovich, 3rd edition.

Laganaro, Marina and François-Xavier Alario. (2006) On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language* 55, 178–196.

Lahiri, Aditi, Letitia Gewirth, and Sheila E. Blumstein. (1984) A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *Journal of the Acoustical Society of America* 76, 391–404.

Lakoff, George. (1970) Global rules. *Language* 46, 627–639.

Lakoff, George. (1971) On generative semantics, in *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, ed. by D. D. Steinberg and L. A. Jakobovits, 232–296, Cambridge: Cambridge University Press.

Lakoff, George. (1973) Some thoughts on transderivational constraints, in *Issues in Linguistics: Papers in Honor of Henry and Renee*, ed. by Braj B. Kachru *et al.*, 442–52, Urbana, IL: University of Illinois Press.

Lakoff, George. (1993) Cognitive phonology, in *The Last Phonological Rule*, ed. by John A. Goldsmith, 117–145, Chicago: University of Chicago Press.

Laks, Bernard. (1995) A connectionist account of French syllabification. *Lingua* 95, 56–75.

Laks, Bernard. (2003) Saussure's phonology, in *Take Danish, for Instance: Linguistic Studies in Honour of Hans Basbøll*, ed. by H. Galberg Jacobsen, D. Bleses, T. O. Madsen, and P. Thomsen, 199–211, Odense: University of Southern Denmark University Press.

Lamontagne, Gregory. (1993) *Syllabification and Consonant Cooccurrence Conditions*. University of Massachusetts, MA: Amherst dissertation.

Lamoureux, Annie-Karine. (2000) Les emprunts lexicaux français en anglais canadien: Adaptations phonologiques,

importations, et influence orthographique. Master's thesis, Laval University, Quebec City.
Laniran, Yetundo O. and G. N. Clements. (2003) Downstep and high raising: Interacting factors in Yoruba tone production. *Journal of Phonetics* 31, 203–250.
Larson, Gary. (1992) *Dyamic Computational Models and the Representation of Phonological Information*. University of Chicago dissertation.
Lass, Roger. (1976) *English Phonology and Phonological Theory: Synchronic and Diachronic Studies*. Cambridge University Press.
Lass, Roger. (1984) *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.
Lass, Roger. (1997) *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
Laver, John. (1994) *Principles of Phonetics*. Cambridge: Cambridge University Press.
Law, Sam-po. (1990) *The Syntax and Phonology of Cantonese Sentence-final Particles*. Boston University dissertation.
Laycock, Don. (1972) Towards a typology of ludlings, or play languages, in *Linguistic Communications (Working Papers of the Linguistic Society of Australia)*, volume 6, 61–113, Linguistic Society of Australia.
Leben, William R. (1973a) The role of tone in segmental phonology, in *Consonant Types and, Tone*, ed. by Larry M. Hyman, 117–149, Los Angeles: University of Southern California.
Leben, William R. (1973b) *Suprasegmental Phonology*. Cambridge, MA: MIT dissertation.
Leben, William R. (1976) The tones in English intonation. *Linguistic Analysis* 2, 69–107.
Leben, William R. (1978) The representation of tone, in *Tone: A Linguistic Survey*, ed. by Victoria A. Fromkin, 177–219, New York: Academic Press.
Leben, William R. (1980) A metrical analysis of length. *Linguistic Inquiry* 11, 497–509.
Lee, Chung-min, Matthew Gordon and Daniel Büring. (2007) *Topic and Focus: Crosslinguistic Perspectives on Intonation and Meaning*. Dordrecht: Springer.
Lee, Lin-shan, Chiuyu Tseng, and Chingjiang Hsieh. (1993) Improved tone concatenation rules in a formant-based Chinese text-to-speech system. *IEEE Transactions on Speech and Audio Processing* 1, 287–294.
Lee, Ok Joo. (2000) The pragmatics and intonation of *ma*-particle questions in Mandarin. Master's thesis, Ohio State University.
Lee, Ok Joo. (2005) *The Prosody of Questions in Beijing Mandarin*. Ohio State University dissertation.
Lee, Yongeun. (2006) *Sub-Syllabic Constituency in Korean and English*. Northwestern University dissertation.
Lee, Yongeun and Matt Goldrick. (2008) The emergence of sub-syllabic representations. *Journal of Memory and Language* 59, 155–168.
Leer, Jeff. (1985) Prosody in Alutiiq, in *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies*, ed. by Michael Krauss, 77–134, Fairbanks, AK: Alaska Native Language Center.
Legendre, Geraldine, Yoshiro Miyata, and Paul Smolensky. (1990a) Harmonic grammar: A formal multilevel connectionist theory of linguistic well-formedness: Theoretical foundations. Technical Report Technical Report 90–95, Institute of Cognitive Science, University of Colorado.
Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. (1990b) Can connectionism contribute to syntax? Harmonic grammar, with an application, in *Papers from the 26th Annual Meeting of the Chicago Linguistic Society*, ed. by Michael Ziolkowski, Manuela Noske, and Karen Deaton, 237–252, Chicago: Chicago Linguistic Society.

Legendre, Géraldine, Antonella Sorace, and Paul Smolensky. (2006) The Optimality Theory-harmonic grammar connection, in *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar, Volume 2: Linguistic and Philosophical Implications*, ed. by Paul Smolensky and Géraldine Legendre, 339–402, Cambridge, MA: MIT Press/Bradford Books.

Lehiste, Ilse. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, Ilse. (1973) Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America* 54, 1228–1234.

Lehiste, Ilse and Pavle Ivić. (1963) *Accent in SerboCroatian*. Ann Arbor, MI: University of Michigan, Department of Slavic Languages and Literatures.

Lehn, Walter. (1963) Emphasis in Cairo Arabic. *Language* 39, 29–39.

Leopold, W. (1939–1949) *Speech Development of a Bilingual Child: A Linguist's Record*, volumes 1–4. Evanston IL: Northwestern University Press.

Leslau, Wolf. (1995) *A Reference Grammar of Amharic*. Wiesbaden: Harrassowitz.

Levelt, C. (1995) Segmental structure of early words: Articulatory frames or phonological constraints, in *The Proceedings of the Twenty-seventh Annual Child Language Research Forum*, 19–27, Stanford, CA: CSLI.

Levelt, C. C., N. O. Schiller, and W. J. Levelt. (2000) The acquisition of syllabic types. *Language Acquisition* 8, 237–264.

Levelt, Willem J. M. (1974) *Formal Grammars in Linguistics and Psycholinguistics, Vol. III: Psycholinguistic Applications*. The Hague: Benjamins.

Levelt, Willem J. M. (1976) Formal grammars and the natural language user: A review, in *Topics in Artificial Intelligence*, ed. by A. Marzollo, Berlin: Springer-Verlag.

Levelt, Willem J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levelt, Willem J. M., Ardi Roelofs, and Antje S. Meyer. (1999) A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–38.

Levi, Susannah. (2001) Glides, laterals, and Turkish vowel harmony, in *Papers from the 37th Annual Meeting of the Chicago Linguistic Society*, 379–94, Chicago: Chicago Linguistic Society.

Levi, Susannah. (2004) *The Representation of Underlying Glides: A Crosslinguistic Study*. University of Washington dissertation.

Levi, Susannah. (2008) Reconsidering the variable status of glottals in nasal harmony, in *Papers from the 41th Annual Meeting of the Chicago Linguistic Society*, 299–312, Chicago: Chicago Linguistic Society.

Levin, Juliette. (1985) *A Metrical Theory of Syllabicity*. MIT dissertation.

Levine, Robert and Detmar Meurers. (2006) Declarative models of syntax, in *Encyclopedia of Language and Linguistics*, ed. by Keith Brown, volume 3, 377–384, Oxford: Elsevier, 2nd edition.

Levitt, Andrea G. and Alice F. Healy. (1985) The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language* 24, 717–733.

Lewis, Geoffrey. (1967) *Turkish Grammar*. Oxford: Oxford University Press.

Li, Bing. (1996) *Tungusic Vowel Harmony: Description and Analysis*. The Hague: Holland Academic Graphics.

Li, Zhiqiang. (2003) *The Phonetics and Phonology of Tone Mapping in a Constraint-Based Approach*. MIT dissertation.

Liberman, Mark. (1975) *The Intonational System of English*. MIT dissertation.

Liberman, Mark and Janet Pierrehumbert. (1984) Intonational invariance under changes in pitch range and length, in *Language Sound Structure*, ed. by Mark Aronoff and Richard T. Oehrle, 157–233, Cambridge, MA: MIT Press.

Liberman, Mark and Alan S. Prince. (1977) On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–336.

Liddell, Scott. (1978) Nonmanual signals and relative clauses in American Sign Language, in *Understanding Language through Sign Language Research*, ed. by P. Siple, 59–90, New York: Academic Press.

Liddell, Scott. (1984) THINK and BELIEVE: Sequentiality in American Sign Language. *Language* 60, 372–392.

Liddell, Scott and Robert Johnson. (1989) American Sign Language: The phonological base. *Sign Language Studies* 64, 197–277.

Lieber, Rochelle. (1987) *An Integrated Theory of Autosegmental Processes*. Albany, NY: State University of New York Press.

Lieberman, Philip. (1963) Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech* 6, 172–187.

Lieberman, Philip. (1967) *Intonation, Perception, and Language*. Cambridge, MA: MIT Press.

Lightfoot, David. (1999) *The Development of Language*. Cambridge, MA: MIT Press.

Lightner, Theodore. (1965) *Segmental Phonology of Modern Standard Russian*. MIT dissertation.

Lightner, Theodore M. (1975) The role of derivational morphology in generative grammar. *Language* 51, 617–638.

Lim, Laureen T. and Gregory R. Guy. (2005) The limits of linguistic community: speech styles and variable constraint effects, in *Penn Working Papers in Linguistics*, vol. 10.2. Selected papers from NWAVE 32, ed. by Maciej Baranowski, Uri Horest, Keelan Evans, and Giang Nguyen, 157–170, Philadelphia: Penn Linguistics Club.

Lin, Ying. (2005a) *Learning Features and Segments from Waveforms: A Statistical Model of Early Phonological Acquisition*. UCLA dissertation.

Lin, Ying. (2005b) *Learning Features and Segments from Waveforms: A Statistical Model of Early Phonological Acquisition*, UCLA dissertation.

Lin, Ying. (2005c) Learning stochastic OT grammars: A Bayesian approach using data augmentation and Gibbs sampling, in *Proceedings of the Association for Computational Linguistics*, East Stroudsburg, PA: Association for Computational Linguistics.

Lin, Ying. (2006) Stochastic Optimality Theory, local search, and Bayesian learning of hierarchical linguistic models. Department of Linguistics, University of Arizona.

Lindblom, Björn. (1983) Economy of speech gestures, in *Speech Production*, ed. by Peter F. MacNeilage, 217–245, New York: Springer-Verlag.

Lindblom, Björn. (1990) Explaining phonetic variation: A sketch of the H & H theory, in *Speech Production and Speech Modelling*, ed. by William Hardcastle and Alain Marchal, 403–439, Dordrecht: Kluwer Academic Press.

Lindblom, Björn. (1992) Phonological units as adaptive emergents of lexical development, in *Phonological Development: Models, Research, Implications*, ed. by Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon, Timonium, MD: York Press.

Lindblom, Björn, Susan G. Guion, Susan Hura, Seung-Jae Moon, and Raquel Willerman. (1995) Is sound change adaptive? *Rivista di Linguistica* 7, 5–36.

Lindblom, Björn. (1986) Phonetic universals in vowel systems, in *Experimental Phonology*, ed. by J. Ohala and J. Jaeger, 13–44, Orlando: Academic Press.

Linebaugh, Gary. (2007) *Phonetic Grounding and Phonology: Vowel Backness Harmony and Vowel Height Harmony*. University of Illinois, Urbana-Champaign dissertation.

Linell, Per. (1979) *Psychological Reality in Phonology: A Theoretical Study*. Cambridge: Cambridge University Press.

Ling, Charles X. and Marin Marinov. (1993) Answering the connectionist

challenge: A symbolic model of learning the past tenses of English verbs. *Cognition* 49, 235–290.
Liphola, Marcelino M. (1999) Neutralizing vowel reduction in Shimakonde. Manuscript, Ohio State University.
Liphola, Marcelino M. (2001) *Aspects of the Phonology and Morphology of Shimakonda*. Ohio State University dissertation.
Lleó, Conxiita (1996) To spread or not to spread: Different styles in the acquisition of Spanish phonology, in *Proceedings of the UBC International Conference on Phonological Acquisition*, ed. by B. H. Bernhardt, J. Gilbert, and D. Ingram, 215–228, Somerville, MA: Cascadilla Press.
Lleó, Conxiita (2001) The interface of phonology and morphology: The emergence of the article in the early acquisition of Spanish and German, in *Approaches to Bootstrapping: Phonological, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, ed. by J. Weissenborn and B. Höhle, 23–44, Amsterdam: Benjamins.
Lleó, Conxiita (2006) The acquisition of prosodic word structures in Spanish by monolingual and Spanish-German bilingual children. *Language and Speech* 49, 207–231.
Lleó, Conxiita and K. Demuth. (1999) Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and Romance languages, in *Proceedings of the 23rd Annual Boston University Conference on Language Development*, ed. by A. Greenhill, H. Littlefield, and C. Tano, 407–418, Somerville, MA: Cascadilla Press.
Lleó, Conxiita and M. Prinz. (1996) Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language* 23, 31–56.
Lloyd, R. and J. Lloyd. (1992) Baruya organised phonology data. Available at SIL Papua New Guinea web site.
Local, John. (2007) Phonetic detail and the organization of talk-in-interaction, in *Proceedings of the 16th International Congress of Phonetic Sciences*, 1–10, Saarbrüecken: Universität des Saarlandes.
Locke, John L. (1983) *Phonological Acquisition and Change*. New York: Academic Press.
Locke, John L. (1995) *The Child's Path to Spoken Language*. Cambridge, MA: Harvard University Press.
Loehr, Daniel. (2004) *Gesture and Intonation*. Georgetown University, dissertation.
Lombardi, Linda. (1991) *Laryngeal Features and Laryngeal Neutralization*. University of Masschusetts, Amherst dissertation.
Lombardi, Linda. (1999) Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17, 267–302.
Lombardi, Linda. (2001) Why place and voice are different: Constraint-specific alternations in Optimality Theory, in *Segmental phonology in Optimality Theory: Constraints and Representations*, ed. by Linda Lombardi, 13–45, Cambridge: Cambridge University Press.
Longacre, Robert E. (1957) *Proto-Mixtecan*. Bloomington, IN: Indiana University.
Loos, Eugene. (1969) *The Phonology of Capanahua and its Grammatical Basis*, volume 20. University of Oklahoma: Summer Institute of Linguistics.
Lovins, Julie. (1975) *Loanwords and Phonological Structure of Japanese*. Bloomington, IN: Indiana University Linguistics Club.
Lowenstamm, Jean. (1996) CV as the only syllable type, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks, volume 2, 419–442, Manchester: European Studies Research Institute.
Lowenstamm, Jean. (1999) The beginning of the word, in *Phonologica 1996: Syllables!?*, ed. by John Rennison

and Klaus Kühnhammer, 153–166, The Hague: Thesus.
Lowenstamm, Jean. (2003) Remarks on *mutæ cum liquidā* and branching onsets, in *Living on the Edge: 28 Papers in Honor of Jonathan Kaye*, ed. by Stefan Ploch, volume 62, 339–363, Berlin and New York: Mouton de Gruyter.
Łubowicz, Anna. (2002) Derived environment effects in optimality theory. *Lingua* 112, 243–280.
Łubowicz, Anna. (2003a) *Contrast Preservation in Phonological Mappings*. University of Massachusetts, Amherst dissertation.
Łubowicz, Anna. (2003b) Countering-feeding opacity as a chain shift effect, in *Papers from the 22nd West Coast Conference on Formal Linguistics*, 315–327.
Łubowicz, Anna. (2007) Paradigmatic contrast in Polish. *Journal of Slavic Linguistics* 15, 229–262.
Łubowicz, Anna. (forthcoming) Chain shifts, in *The Blackwell Companion to Phonology*, ed. by Marc van Oostendorp, Colin Ewen, Elizabeth V. Hume, and Keren Rice, Oxford: Wiley-Blackwell.
Luce, Paul A., Stephen D. Goldinger, Jr. Edward T. Auer and Michael S. Vitevitch. (2000) Phonetic priming, neighborhood activation and PARSYN. *Perception and Psychophysics* 62, 615–625.
Lugt, Arie H. van der. (2001) The use of sequential probabilities in the segmentation of speech. *Perception and Psychophysics* 63, 811–823.
Lunt, Horace. (1952) *A Grammar of the Macedonian Literary Language*. Skopje.
Luria, A. and F. Yudovich. (1959) *Speech and Development of Mental Processes in the Child*. London: Staples.
Lyberg, Bertil. (1981) Some consequences of a model for segment duration based on F0 dependence. *Journal of Phonetics* 9, 97–103.
MacBride, Alexander. (2004) *A Constraint-Based Approach, to Morphology*. UCLA dissertation.
MacDonald, George and Georgetta MacDonald. (1974) Phonologies of four Papua New Guinea languages, in *Workpapers in Papua New Guinea Languages*, ed. by Richard Loving, volume 4, 129–158, Summer Institute of Linguistics.
MacDonald, Lorna. (1990) *A Grammar of Tauya*. New York: Mouton de Gruyter.
MacDonald, M. C., N. J. Pearlmutter, and M. S. Seidenberg. (1994) Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101, 676–703.
MacEachern, Margaret R. (1997) *Laryngeal Cooccurrence Restrictions*. UCLA dissertation. Published by Garland Publishing, New York, 1999.
MacKay, Donald G. (1972) The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology* 3, 210–227.
Macken, Marlys. (1978) Permitted complexity in phonological development: One child's acquisition of Spanish consonants. *Lingua* 44, 219–253.
Macken, Marlys. (1979) Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua* 49, 11–49.
Macken, Marlys. (1980) The child's lexical representation: The "puzzle-puddle-pickle" evidence. *Journal of Linguistics* 16, 1–17.
Macken, Marlys. (1987) Representation, rules, and overgeneralization in phonology, in *Mechanisms of Language Acquisition*, ed. by Brian MacWhinney, 367–397, Hillsdale, NJ: Lawrence Erlbaum Associates.
Macken, Marlys. (1995) Phonological acquisition, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 671–696, Oxford: Blackwell.
Macken, Marlys and David Barton. (1980a) The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. *Journal of Child Language* 7, 433–458.

Macken, Marlys and David Barton. (1980b) A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language* 7, 41–74.

Macken, Marlys and Charles A. Ferguson. (1981) Phonological universals in language acquisition, in *Native Language and Foreign Language Acquisition*, ed. by H. Winitz, 110–129, New York: New York Academy of Sciences.

Macken, Marlys and Charles A. Ferguson. (1983) Cognitive aspects of phonological development: Model, evidence and issues, in *Children's Language*, ed. by K. E. Nelson, volume 4, 255–282, Hillsdale, NJ: Lawrence Erlbaum.

Mackenzie, Sara. (2005) Similarity and contrast in consonant harmony systems, in *Toronto Working Papers in Linguistics*, volume 24, 169–182, University of Toronto.

Mackey, William F. (1992) La génèse d'une typologie de diglossie. *Revue de langues et linguistique*, 31–43.

MacNamara, John. (1984) *Names for Things: A Study of Human Learning*. Cambridge, MA: MIT Press.

MacNeilage, Peter F. (1980) The control of speech production, in *Child Phonology*, ed. by G. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson, volume 1, 9–21, New York: Academic Press.

MacNeilage, Peter F. and Barbara L. Davis. (1990) Acquisition of speech production: The achievement of segmental independence, in *Speech Production and Speech Modeling*, ed. by W. I. Hardcastle and A. Marchal, 55–68, Kluwer: Dordrecht.

MacNeilage, Peter F. and Barbara L. Davis. (2000) Deriving speech from nonspeech: A view from ontogeny. *Phonetica* 57, 284–296.

MacWhinney, Brian. (1975) Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language* 2, 65–77.

MacWhinney, Brian. (2000) *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maddieson, Ian. (1985) Phonetic cues to syllabification, in *Phonetic Linguistics*, ed. by Victoria A. Fromkin, 203–221, Academic Press.

Maddieson, Ian. (1997) Phonetic universals, in *The Handbook of Phonetic Sciences*, ed. by William Hardcastle and John Laver, 619–639, Oxford: Blackwell.

Maddieson, Ian and Kristin Precoda. (1990) Updating UPSID. *UCLA Working Papers in Phonetics* 74, 104–111.

Maekawa, Kikuo. (2003) Corpus of spontaneous Japanese: Its design and evaluation, in *Proceedings of the ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003), Tokyo*, 7–12.

Maekawa, Kikuo and Hideaki Kikuchi. (2005) Corpus-based analysis of vowel devoicing in spontaneous Japanese: An interim report, in *Voicing in Japanese*, ed. by Jeroen van de Weijer, Kensuke Najo, and Tetsuo Nishihara, 205–228, Berlin: Mouton de Gruyter.

Maekawa, Kikuo, Hideaki Kikuchi, Yosuke Igarashi and Jennifer J. Venditti. (2002) X-JToBI: An extended J-ToBI for spontaneous speech, in *Proceedings of the 7th International Conference on Spoken Language Processing*, Boulder, CO, 1545–1548.

Magnuson, James S., Bob McMurray, Michael K. Tanenhaus and Richard N. Aslin. (2003a) Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science* 27, 801–805.

Magnuson, James S., Bob McMurray, Michael K. Tanenhaus, and Richard N. Aslin. (2003b) Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science* 27, 285–298.

Magri, Giorgio (2010) HG has no computational advantages over OT:

consequences for the theory of OT online algorithms. Rutgers Optimality Archive ROA-1104, http://roa.rutgers.edu/view.php3?id=1600 (Accessed April 23, 2011).

Mahanta, Shakuntala. (2007) *Directionality and Locality in Vowel Harmony*. Utrecht University dissertation.

Maiden, Martin. (2005) Morphological autonomy and diachrony, in *Yearbook of Morphology 2004*, ed. by Geert Booij and Jaap van Marle, 137–175, Berlin: Springer.

Majerus, Steve and Martial Van der Linden. (2003) Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology* 21, 303–310.

Majerus, Steve, Martial Van der Linden, Ludivine Mulder, Thierry Meulemans, and Frédéric Peters. (2004) Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language* 51, 297–306.

Majors, Tivoli. (1998) *Stress-dependent Harmony*. University of Texas, Austin dissertation.

Mak, Cornelia. (1953) A comparison of two Mixtec tonemic systems. *International Journal of American Linguistics* 19, 85–100.

Makkai, Valeria Becker (ed.) (1972) *Phonological Theory: Evolution and Current Practice*. New York: Holt, Rinehart, and Winston.

Malmstrom, P. M. and M. N. Silva. (1986) Twin talk: manifestations of twins status in the speech of toddlers. *Journal of Child Language* 13, 293–304.

Malécot, André. (1958) The role of releases in identification of released final stops. *Language* 34, 370–380.

Manaster-Ramer, Alexis. (1996) A letter from an incompletely neutral phonologist. *Journal of Phonetics* 24, 477–489.

Mandel, Mark A. (1981) *Phonotactics and Morphophonology in American Sign Language.* University of California, Berkeley dissertation.

Mandelbrot, B. (1961) On the theory of word frequencies and on related Markovian models of discourse, in *Proceedings of the 12th Symposium in Applied Mathematics Volume XII: Structure of Language and its Mathematical Aspects*, ed. by R. Jakobson, 190–219, Providence, RI: American Mathematical Society.

Manuel, Sharon. (1999) Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts, in *Coarticulation: Theory, Data and Techniques*, ed. by W. J. Hardcastle and N. Hewitt, 179–198, Cambridge: Cambridge University Press.

Manuel, Sharon Y. and Rena A. Krakow. (1984) Universal and language particular aspects of vowel-to-vowel coarticulation. Status Reports on Speech Research SR-77/78, 69-78, Haskins Laboratories. Haskins Laboratories Status Report.

Marantz, Alec. (1982) *Re* reduplication. *Linguistic Inquiry* 18, 483–545.

Markov, Andrey A. (1913) Essai d'une recherche statistique sur le texte du roman "Eugene Onegin" illustrant la liaison des épreuves en chaîne. *Bulletin de l'Academie Imperiale des Science de St. Petersbourg* 7, 153–169.

Marr, David. (1982) *Vision.* San Francisco, CA: W. H. Freeman and Company.

Martin, Andrew. (2006) The correlation of markedness and frequency: Evidence from Latin and French, in *Proceedings of the 37th Annual Meeting of the North East Linguistic Society*, ed. by Emily Elfner and Martin Walkow, volume 37, Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts.

Martin, Andrew. (2007) *The Evolving Lexicon.* UCLA dissertation.

Martin, Jack and Margaret McKane Mauldin. (2000) *A Dictionary of Creek/*

Muskogee, With Notes on the Florida and Oklahoma Seminole Dialects of Creek. Lincoln, NE: University of Nebraska Press.

Martin, Nadine and Prahlad Gupta. (2004) Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology* 21, 213–228.

Martin, Nadine and Eleanor M. Saffran. (2002) The relationship of input and output phonological processing: An evaluation of models and evidence to support them. *Aphasiology* 16, 107–150.

Martin, Pierre. (2002) Le système vocalique du français du Québec: De l'acoustique à la phonologie. *La Linguistique* 38, 71–88.

Martinet, André. (1939) Un ou deux phonèmes? *Acta Linguistics* 1, 94–103.

Martinet, André. (1952) Function, structure, and sound change. *Word* 8, 1–32.

Martinet, André. (1955) *Économie des changements phonétiques: traité de phonologie diachronique.* Bern: Francke.

Martinet, André. (1964) *Elements of General Linguistics.* Chicago: University of Chicago Press.

Mascaró, Joan. (1976) *Catalan Phonology and the Phonological Cycle.* MIT dissertation.

Maslova, Elena. (2005) Stochastic OT as a model of constraint interaction, in *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*, ed. by Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson, and Annie Zaenen, Stanford, CA: CSLI Publications.

Massaro, Dominic W. and Michael M. Cohen. (1983) Phonological context in speech perception. *Perception and Psychophysics* 34, 338–348.

Mathur, Gaurav and Christian Rathmann. (2006) Variability in verbal agreement forms across four sign languages, in *Papers from Laboratory Phonology VIII: Varieties of Phonological Competence*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 285–314, The Hague, Mouton.

Mathur, Gaurav and Christian Rathmann. (2010) Verb agreement in sign language morphology, in *Sign Language: A Cambridge Language Survey*, ed. by Diane Brentari, 173–196, Cambridge: Cambridge University Press.

Matisoff, James A. (1973) Tonogenesis in Southeast Asia, in *Consonant Types and Tone*, ed. by Larry M. Hyman, of *Southern California Occasional Papers in Linguistics*, volume 1, 71–95, Los Angeles: University of Southern California.

Matisoff, James A. (1975) Rhinoglottophilia: The mysterious connection between nasality and glottality, in *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, ed. by Charles A. Ferguson, Larry M. Hyman, and John J. Ohala, 265–287, Stanford, CA: Language Universals Project, Department of Linguistics, Stanford University.

Matteson, Esther. (1965) *The Piro (Arawakan) Language.* Berkeley, CA: University of California Press.

Matthei, E. H. (1989) Crossing boundaries: More evidence for phonological constraints on early multi-word utterances. *Journal of Child Language* 16, 41–54.

Matthews, Peter H. (1972) *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation.* Cambridge: Cambridge University Press.

Mattys, Sven L. and Julie M. Liss. (2008) On building models of spoken-word recognition: When there is as much to learn from natural "oddities" as artificial normality. *Perception and Psychophysics* 70, 1235–1242.

Maxwell, Michael. (1994) Parsing using linearly ordered phonological rules, in *Computational Phonology: First Meeting of the ACL Special Interest Group in Computational Phonology*, ed. by S. Bird,

59–70, Association for Computational Linguistics.
May, Jean and Eunice Loeweke. (1964) The phonological hierarchy in Fasu. *Anthropological Linguistics* 7, 89–97.
Maye, Jessica, Janet F. Werker, and Louann Gerken. (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111.
McCarthy, John J. (1979a) *Formal Problems in Semitic Phonology and Morphology.* MIT dissertation.
McCarthy, John J. (1979b) On stress and syllabification. *Linguistic Inquiry* 10, 443–466.
McCarthy, John J. (1981) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–418.
McCarthy, John J. (1982) Prosodic structure and expletive infixation. *Language* 58, 574–590.
McCarthy, John J. (1986) OCP effects: Gemination and antigemination. *Linguistic Inquiry* 20, 71–99.
McCarthy, John J. (1988a) Feature geometry and dependency: A review. *Phonetica* 43, 84–108.
McCarthy, John J. (1988b) Feature geometry and dependency: A review. *Phonetica* 43, 84–108.
McCarthy, John J. (1989) Linear order in phonological representation. *Linguistic Inquiry* 20, 71–99.
McCarthy, John J. (1991) L'infixation réduplicative dans les langages secrets. *Langages* 101, 11–29.
McCarthy, John J. (1994a) On coronal "transparency." Presented at the Trilateral Phonology Weekend.
McCarthy, John J. (1994b) The phonetics and phonology of Semitic pharyngeals, in *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, ed. by Patricia Keating, 191–233, Cambridge: Cambridge University Press.
McCarthy, John J. (1997) Process-specific constraints in Optimality Theory. *Linguistic Inquiry* 28, 231–251.
McCarthy, John J. (1998) Morpheme structure constraints and paradigm occultation, in *Papers from the 32th Annual Meeting of the Chicago Linguistic Society, Vol. II: The Panels*, ed. by M. Catherine Gruber, Derrick Higgins, Ken Olson, and Tamra Wysocki, volume 2, Chicago: Chicago Linguistic Society.
McCarthy, John J. (2002) *A Thematic Guide to Optimality Theory.* Cambridge: Cambridge University Press.
McCarthy, John J. (2003a) Comparative markedness. *Theoretical Linguistics* 29, 1–51.
McCarthy, John J. (2003b) OT constraints are categorical. *Phonology* 20, 75–138.
McCarthy, John J. (2003c) Sympathy, cumulativity, and the Duke-of-York gambit, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Jijver, Cambridge: Cambridge University Press.
McCarthy, John J. (2004) Headed spans and autosegmental spreading. Manuscript, University of Massachusetts, Amherst.
McCarthy, John J. (2005) Taking a free ride in morphononemic learning. *Catalan Journal of Linguistics* 4, 19–55. ROA 683.
McCarthy, John J. (2007a) Consonant harmony via correspondence: Evidence from Chumash, in *Papers in Optimality Theory 3*, ed. by Leah Bateman, Michael O'Keefe, Ehren Reilly, and Adam Werle, 223–237, Amherst, MA: GLSA.
McCarthy, John J. (2007b) Derivations and levels of representation, in *The Cambridge Handbook of Phonology*, ed. by Paul De Lacy, 99–117, Cambridge: Cambridge University Press.
McCarthy, John J. (2007c) *Hidden Generalizations: Phonological Opacity in Optimality Theory.* London: Equinox.
McCarthy, John J. (2007d) What is Optimality Theory? *Language and Linguistics Compass* 1, 260–291.
McCarthy, John J. (2008) *Doing Optimality Theory*. Oxford: Blackwell.

McCarthy, John J. and Alan Prince. (1986) Prosodic morphology. Rutgers Technical Report TR-32.

McCarthy, John J. and Alan Prince. (1993a) Generalized alignment, in *Yearbook of Morphology 1993*, ed. by Geert Booij and Jaap van Marle, 79–153, Dordrecht: Kluwer.

McCarthy, John J. and Alan Prince. (1993b) Prosodic morphology I: Constraint interaction and satisfaction. ROA 48-1201.

McCarthy, John J. and Alan Prince. (1994a) The emergence of the unmarked: Optimality in prosodic morphology, in *Proceedings of the 24th North East Linguistics Society*, ed. by Mercé González, 333–379, Amherst, MA: GLSA.

McCarthy, John J. and Alan Prince. (1994b) An overview of prosodic morphology. Part I: Templatic form in reduplication; Part II: Template satisfaction. Presented at Workshop on Prosodic Morphology, Utrecht University.

McCarthy, John J. and Alan Prince. (1995) Faithfulness and reduplicative identity, in *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. by Jill Beckman, Laura Dickey, and Suzanne Urbanczyk, 249–384, Amherst, MA: GLSA.

McCarthy, John J. and Alan Prince. (1996) Prosodic morphology. Technical Report 32, Rutgers University Center for Cognitive Science. University of Massachusetts, Amherst and Brandeis University, manuscript (1986, unannotated).

McCarthy, John J. and Alan Prince. (1999a) Faithfulness and identity in prosodic morphology, in *The Prosody-Morphology Interface*, ed. by Harry van der Hulst, René Kager, and Wim Zonneveld, 218–309, Cambridge: Cambridge University Press.

McCarthy, John J. and Alan Prince. (1999b) Sympathy and phonological opacity. *Phonology* 16, 331–399.

McCarthy, John J. and Matthew Wolf. (2009) Less than zero: Correspondence and the null output, in *Modeling Ungrammaticality in Optimality Theory*, ed. by Sylvia Blaho and Curt Rice, London: Equinox Publishing.

McCawley, James D. (1963) Consonant mutation in Finnish. *Quarterly Progresss Report, Research Laboratory of Electronics, MIT* 70, 268–273.

McCawley, James D. (1965) *The Accentual System of Standard Japanese.* MIT dissertation.

McCawley, James D. (1967) Sapir's phonologic representation. *International Journal of American Linguistics* 33, 106–111.

McCawley, James D. (1968a) Concerning the base component of a transformational-generative grammar. *Foundations of Language* 4, 243–269.

McCawley, James D. (1968b) *The Phonological Component of a Grammar of Japanese.* The Hague: Mouton.

McCawley, James D. (1970) Some tonal systems that come close to being pich accent but don't quite make it, in *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 526–532.

McCawley, James D. (1973) On the role of notation in generative phonology, in *The Formal Analysis of Natural Language*, ed. by Maurice Gross *et al.*, 51–62, The Hague: The Hague.

McCawley, James D. (1978) What is a tone language?, in *Tone: A Linguistic Survey*, ed. by Victoria A. Fromkin, 113–131, New York: Academic Press.

McClelland James and Jeffrey Elman. (1986) The TRACE model of speech perception. *Cognitive Psychology* 18, 1–86.

McClelland, James, David E. Rumelhart and the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 2. Cambridge, MA: Bradford Books/MIT Press.

McClelland, James L., Daniel Mirman, and Lori L. Holt. (2006) Are there interactive processes in speech perception? *Trends in Cognitive Sciences* 10, 363–369.

McCormick, Susan M. (1982) *Vowel Harmony and Umlaut: Implications for*

a Typology of Accent. Cornell University dissertation.
McCullough, Stephen, Karen Emmorey and Martin Sereno. (2005) Neural organization for recognition of grammatical and emotional facial expressions in deaf ASL signers and hearing Nonsigners. *Cognitive Brain Research* 22, 193–203.
McCune, Lorraine and Marilyn M. Vihman. (1987) Vocal motor schemes. *Papers and Reports on Child Language Development* 26, 72–79.
McKaughan, Howard P. (1954) Chatino formulas and phonemes. *International Journal of American Linguistics* 19, 23–27.
McLennan, Conor T. and Paul A. Luce. (2005) Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning Memory and Cognition* 31, 306–321.
McLennan, Conor T., Paul A. Luce, and Jan Charles-Luce. (2003) Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29, 539–553.
McLeod, Peter, Kim Plunkett, and Edmund T. Rolls. (1998) *Introduction to Connectionist Modelling of Cognitive Processes.* Oxford: Oxford University Press.
McMahon, April. (1994) *Understanding Language Change.* Cambridge: Cambridge University Press.
McMillan, Corey T., Martin Corley, and Robin J. Lickley. (2009) Articulatory evidence for feedback and competition in speech production. *Language and Cognitive Processes* 24, 44–66.
McMurray, Bob, Richard Aslin, and Joseph C. Toscano. (2009) Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Sciences* 12, 369–378.
McNeill, David. (2005) *Language and Gesture*. Cambridge and New York: Cambridge University Press.
McQueen, James. (1998) Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39, 21–46.
McQueen, James. (2003) The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus, and Aslin (2003). *Cognitive Science* 27, 795–799.
McQueen, James, Anne Cutler, and Dennis Norris. (2006) Phonological abstraction in the mental lexicon. *Cognitive Science* 30, 1113–1126.
Medin, Douglas L. and Marguerite M. Schaffer. (1978) Context theory of classification learning. *Psychological Review* 85, 207–238.
Meier, Richard. (2002) Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech, in *Modality in Langauge and Linguistic Theory*, ed. by Richard Meier, K. Cormier, and D. Quinto-Pozos, 1–25, Cambridge: Cambridge University Press.
Meir, Irit, Carol Padden, Mark Aronoff and Wendy Sandler. (2007) Body as subject. *Journal of Linguistics* 43, 531–563.
Melinger, Alissa. (2002) Foot structure and accent in Seneca. *International Journal of American Linguistics* 68, 287–315.
Mellander, Evan. (2001) Quantitative processes in trochaic systems, in *West Coast Conference on Formal Linguistics 20*, 414–427.
Melvold, Janis. (1986) Cyclicity and Russian stress, in *Proceedings of the 17th Annual Meeting of the North East Linguistic Society*, ed. by J. McDonough and B. Plunkett, 467–481, Amherst, MA: GLSA.
Ménard, L., H. Loevenbruck, and C. Savariaux. (2006) Articulatory and acoustic correlates of contrastive focus in French children and adults, in *Speech Production: Models, Phonetic Processes and Techniques*, ed. by J. Harrington and M. Tabain, 227–251, New York: Psychology Press.
Menn, Lise. (1971) Phonotactic rules in beginning speech. *Lingua* 26, 225–251.

Menn, Lise. (1976) Evidence for an interactionist discovery theory of child phonology, in *Papers and Reports on Language Development*, volume 12, 169–177, Stanford, CA: Stanford University.

Menn, Lise. (1983) Development of articulatory, phonetic, and phonological capabilities, in *Language production*, ed. by B. Butterworth, volume 2, 3–50, London: Academic Press.

Menn, Lise and Brian Macwhinney. (1984) The repeated morph constraint: Toward an explanation. *Language* 60, 519–541.

Menn, Lise and E. Matthei. (1992) The "two-lexicon" account of child phonology: Looking back, looking ahead, in *Phonological Development: Models, Research, Implications*, ed. by Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon, 211–247, Timonium, MD: York Press.

Menovshchikov, G. A. (1975) *Iazyk Naukanskikh Eskimosov.* Leningrad: Nauka.

Menyuk, Paula. (1968) Children's learning and reproduction of grammatical and nongrammatical phonological sequences. *Child Development* 39, 849–859.

Merchant, Nazarre and Bruce Tesar. (2008) Learning underlying forms by searching restricted lexical subspaces, in *Papers from the 41st Annual Meeting of the Chicago Linguistic Society*, volume 2, 33–48, Chicago: Chicago Linguistic Society.

Messer, S. (1967) Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior* 6, 609–613.

Mester, Armin. (1988) Dependent tier ordering and the OCP, in *Features, Segmental Structure, and Harmony Processes*, ed. by Harry van der Hulst and Norval Smith, volume 2, 128–144, Dordrecht: Foris.

Mester, Armin. (1994) The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12, 1–61.

Michelson, Karin. (1988) *A Comparative Study of Lake-Iroquoian Accent*. Dordrecht: Kluwer.

Mielke, Jeff. (2004) *The Emergence of Distinctive Features*. Ohio State University dissertation.

Mielke, Jeff. (2005a) Ambivalence and ambiguity in laterals and nasals. *Phonology* 22, 169–203.

Mielke, Jeff. (2005b) Modeling distinctive feature emergence, in *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. by J. Alderete, C. Han, and A. Kochetov, 281–9, Somerville, MA: Cascadilla Press.

Mielke, Jeff. (2008) *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

Mielke, Jeff, Mike Armstrong, and Elizabeth V. Hume. (2003) Looking through opacity. *Theoretical Linguistics* 29, 123–139.

Mikheev, Andrei. (1997) Automatic rule induction for unknown word guessing. *Computational Linguistics* 23, 405–423.

Miller, Christopher. (1996a) *Phonologie de la langue des signes québécoise: Structure, simultanée et axe temporel*. Montreal: Université du Québec à Montreal dissertation.

Miller, D. Gary. (1975) On constraining global rules in phonology. *Language* 51, 128–132.

Miller, George A. and Noam Chomsky. (1963) Finitary models of language users, in *Handbook of Mathematical Psychology*, ed. by R. D. Luce, R. R. Bush, and E. Galanter, volume 2, pp. 419–491, New York: John Wiley.

Miller, Wick. (1996b) Sketch of Shoshone, a Uto-Aztecan language, in *Handbook of American Indian Languages, vol. 17 (Languages)*, ed. by Ives Goddard, 693–720, Washington: Smithsonian Institute.

Milroy, James and John Harris. (1980) When is a merger not a merger? The MEAT/MATE problem in a present-day

English vernacular. *English World-Wide* 1, 199–210.
Minda, J. P. and J. D. Smith. (2002) Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28, 275–292.
Minkova, Donka. (2004) Philology, linguistics, and the history of /hw/ /w/, in *Studies in the History of the English language II: Unfolding Conversations.*, ed. by Anne Curzan and Kimberly Emmons, 7–46, Berlin: Mouton de Gruyter.
Mintz, Toby and Rachel Walker. (2006) Infants' sensitivity to vowel harmony and its role in word segmentation. Presented at the Annual Meeting of the Linguistic Society of America.
Mitterer, Holger and Leo Blomert. (2003) Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychopysics* 65, 956–969.
Miura, Ichiro and Noriyo Hara. (1995) Production and perception of rhetorical questions in Osaka Japanese. *Journal of Phonetics* 23, 291–303.
Miyaoka, Osahito. (1985) Accentuation in Central Alaskan Yupik, in *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies*, ed. by Michael Krauss, 51–76, Fairbanks, AK: Alaska Native Language Center.
Möbius, Bernd. (1993) *Ein quantitatives Modell der deutschen Intonation: Analyse und Synthese von Grundfrequenzverläfen*. Tübingen: Max Niemeyer Verlag.
Möbius, Bernd, Matthias Pätzold, and Wolfgang Hess. (1993) Analysis and synthesis of German F0 contours by means of Fujisaki's model. *Speech Communication* 13, 53–61.
Mochizuki, M. (1981) The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics* 9, 283–303.
Mogford, K. (1993) Language development in twins, in *Language Development in Exceptional Circumstances*, ed. by D. Bishop and K. Mogford, 80–95, Hove, UK: Lawrence Erlbaum.
Mohanan, K. P. (1982) *Lexical Phonology*. MIT dissertation.
Mohanan, K. P. (1986) *The Theory of Lexical Phonology*. Dordrecht: Reidel.
Mohanan, K. P. (1992) Describing the phonology of non-native varieties of a language. *World Englishes.*
Mohanan, K. P. (1995) The organization of the grammar, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 24–69, Oxford: Blackwell Publishers.
Mohanan, K. P. (2000) The theoretical substance of optimality formalism. *The Linguistic Review* 17, 143–167.
Molino, Jean. (1982) Le nom propre dans la langue. *Langages* 66, 1–20.
Montes-Rodrigues, María Emilia. (1995) *Tonología de la lengua ticuna*. Lenguas Aborigenes de Colombia, Bogotá Universidad de los Andes.
Moon, Grace. (1998) Language games in Korean. Manuscript, Harvard University.
Morais, José, São Luís Castro, Leonor Scliar-Cabral, Régine Kolinsky and Alain Content. (1987) The effects of literacy on the recognition of dichotic words. *The Quarterly Journal of Experimental Psychology Section A* 39, 451–465.
Moravcsik, Edith. (1977) *On Rules of Infixing*. Bloomington, IN: Indiana University Linguistics Club.
Moravcsik, Edith. (2000) Infixation, in *Morphologie Morphology*, ed. by L. Christian Geert Booij and J. Mugdan, 545–552, Berlin: Walter de Gruyter.
Morelli, Frida. (2003) The relative harmony of /s+Stop/ onsets: Obstruent clusters and the sonority sequencing principle, in *The Syllable in Optimality Theory*, 356–371, Cambridge: Cambridge University Press.

Morén, Bruce. (1999) *Distinctiveness, Coercion and Sonority: A Unified Theory of Weight*. University of Maryland dissertation.

Morén, Bruce. (2000) The puzzle of Kashmiri stress. *Phonology* 17, 365–396.

Morén, Bruce. (2007) The division of labor between segment-internal structure and violable constraints, in *Freedom of Analysis?*, 313–344, Berlin: de Gruyter.

Moreton, Elliott. (2002) Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84, 55–71.

Moreton, Elliott. (2006) Phonotactic learning and phonological typology. Presented at the 37th Annual Meeting of the North East Linguistic Society.

Moreton, Elliott. (2008a) Analytic bias and phonological typology. *Phonology* 25, 83–127.

Moreton, Elliott. (2008b) Modelling modularity bias in phonological pattern learning, in *Proceedings of the 27th Annual Meeting of the West Coast Conference on Formal Linguistics*, ed. by Natasha Abner, Jason Bishop, and Kevin Ryan, 1–16, Somerville, MA: Cascadilla Press.

Moreton, Elliott. (2010) Underphonologization and modularity bias, in *Phonological Argumentation: Essays on Evidence and Motivation*, ed. by Steve Parker, 79–101, London: Equinox.

Moreton, Elliott and Shigeaki Amano. (1999) Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies, in *Eurospeech '99*, 2679–2682.

Moreton, Elliott, Gary Feng, and Jennifer Smith. (2005) Syllabification, sonority, and perception: New evidence from a language game, in *Papers from the 41st Annual Meeting of the Chicago Linguistic Society*, 341–355, Chicago: Chicago Linguistic Society.

Moreton, Elliott and Paul Smolensky. (2002) Typological consequences of local constraint conjunction, in *Proceedings of the 21st Annual Meeting of the West Coast Conference on Formal Linguistics*, 306–319.

Morgan, J. L. and K. Demuth. (1996) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Morrisette, M. L., D. A. Dinnsen, and J. A. Gierut. (2003) Markedness and context effects in the acquisition of place features. *Canadian Journal of Linguistics* 48, 329–356.

Mortensen, David. (2006) *Formal and Substantive Scales in Phonology*. University of California, Berkeley dissertation.

Moskowitz, A. I. (1973) The acquisition of phonology and syntax: A preliminary study, in *Approaches to natural language*, ed. by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes, Dordrecht: Reidel.

Motley, M. T. and B. J. Baars. (1975) Encoding sensitivities to phonological markedness and transitional probability: Evidence from spoonerisms. *Human Communication Research* 2, 351–361.

Moulines, Eric and Francis Charpentier. (1990) Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9, 453–467.

Mous, Maarten. (1993) *A Grammar of Iraqw*. Hamburg: Helmut Buske Verlag.

Mowrey, Richard and William Pagliuca. (1995) The reductive character of articulatory evolution. *Rivista di Linguistica* 7, 37–124.

Muller, Jen. (2001) *The Phonology and Phonetics of Word-Initial Geminates*. Ohio State University dissertation.

Müller, Karin. (2002) Probabilistic context-free grammars for phonology, in *Morphological and Phonological Learning: Proceedings of the 6th Workshop of the ACL Special Interest Group in Computational Phonology (SIGPHON)*, 70–80, New Brunswick, NJ: Association for Computational Linguistics.

Munro, Pamela and Charles Ulrich. (1984) Structure-preservation and Western Muskogean rhythmic lengthening, in *West Coast Conference on Formal Linguistics 3*, 191–202.

Munro, Pamela and Catherine Willmond. (1994) *Chickasaw: An Analytical Dictionary*. Norman, OK: University of Oklahoma Press.

Munson, Benjamin. (2001) Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research* 44, 778–792.

Munson, Benjamin and Nancy Pearl Solomon. (2004) The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research* 47, 1048–1058.

Murray, R. and T. Vennemann. (1983) Sound change and syllable structure: Problems in Germanic phonology. *Language* 59, 514–528.

Mutaka, Ngessimo. (1994) *The Lexical Tonology of Kinande*. Munich: Lincom Europa.

Mutaka, Ngessimo. (1995) Vowel harmony in Kinande. *Journal of West African Languages* 25, 41–55.

Mutaka, Ngessimo and Larry M. Hyman. (1990) Syllable and morpheme integrity in Kinande reduplication. *Phonology* 7, 73–120.

Myers, Scott. (1987) *Tone and the Structure of Words in Shona*. University of Massachusetts, Amherst dissertation. Published by Garland Publishing, New York, 1990.

Myers, Scott. (1997) OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15, 847–892.

Myers, Scott. (1998) Surface underspecification of tone in Chichewa. *Phonology* 15, 367–391.

Myers, Scott and Benjamin B. Hansen. (2005) The origin of vowel length neutralisation in vocoid sequences: Evidence from Finnish speakers. *Phonology* 22, 317–344.

Nagahara, Hiroyuki. (1994) *Phonological Phrasing in Japanese*. UCLA dissertation.

Nagy, N. and B. Reynolds. (1997) Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9, 37–55.

Nakatani, Christine Hisayo. (1997) *The Computational Processing of Intonational Prominence: A Functional Prosody Perspective*. Harvard University dissertation.

Nakisa, Ramin C., Kim Plunkett and Ulrike Hahn. (1997) A cross-linguistic comparison of single and dual-route models of inflectional morphology, in *Cognitive Models of Language Acquisition*, ed. by P. Broeder and J. Murre, Cambridge, MA: MIT Press.

Nash, David. (1979) Warlpiri vowel assimilations, in *MIT Working Papers in Linguistics. Vol 1. Papers on Syllable Structure, Metrical Structure and Harmony Processes*, ed. by Ken Safir, 12–24, Cambridge, MA: MIT Press.

Nash, David. (1986) *Topics in Warlpiri Grammar*. New York: Garland.

Nathan, Geoffrey. (1987) On second language acquisition of voiced stops. *Journal of Phonetics* 15, 313–322.

Nazzi, T., G. Iakimova, J. Bertoncini, S. Frédonie, and C. Alcantara. (2006) Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language* 54, 283–299.

Nazzi, T. and F. Ramus. (2003) Perception and acquisition of linguistic rhythm by infants. *Speech Communication* 41, 233–244.

Nearey, Terrance M. (1990) The segment as a unit of speech perception. *Journal of Phonetics* 18, 347–373.

Nearey, Terrance M. and Sherrie E. Shammass. (1987) Formant transitions as partly distinctive invariant properties in the identification of voiced stops. *Canadian Acoustics* 15, 17–24.

Neeleman, Ad and Jan van der Koot. (2006) On syntactic and phonological representations. *Lingua* 116, 1524–1552.
Neidle, Carol, Judy Kegl, Dawn MacLaughlin, Ben Bahan, and Robert G. Lee. (2000) *The Syntax of American Sign Language*. Cambridge, MA: MIT Press.
Nelson, Nicole. (2003) *Asymmetric Anchoring*. Rutgers University dissertation.
Nelson, Nicole. (2004) Wrong-side reduplication is epiphenomenal: Evidence from Yoruba, in *Studies in Reduplication*, ed. by Bernhard Hurch, 135–160, Berlin: Mouton de Gruyter.
Nespor, Marina and Wendy Sandler. (1999) Prosody in Israeli Sign Language. *Language and Speech* 42, 143–176.
Nespor, Marina and Irene Vogel. (1979) Clash avoidance in Italian. *Linguistic Inquiry* 10, 476–482.
Nespor, Marina and Irene Vogel. (1982) Prosodic domains of external sandhi rules, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 1, 224–255, Dordrecht: Foris.
Nespor, Marina and Irene Vogel. (1986) *Prosodic Phonology*. Dordrecht: Foris.
Neu, H. (1980) Ranking of constraints on /t,d/ deletion in American English: A statistical analysis, in *Locating Language in Time and, Space*, ed. by William Labov, 37–54, New York: Academic Press.
Nevins, Andrew. (2004) *Conditions on (Dis)Harmony*. MIT dissertation.
Nevins, Andrew. (2005) Overwriting does not optimize in nonconcatenative word formation. *Linguistic Inquiry* 36, 275–287.
Nevins, Andrew and Ansgar Endress. (2007) The edge of order: Analytic bias in ludlings, in *Harvard Working Papers in Linguistics*, ed. by Jeremy Rau, Keith Plaster, Patrick Liu, and Yaroslav Gorbachov, volume 12, 43–53.
Nevins, Andrew and J.-M. Toro. (2007) Consonants as skeleta, vowels as glue. Presented at the Ljubljana Linguistics Circle.
Nevins, Andrew and Bert Vaux. (2003) Underdetermination in language games: Survey and analysis of Pig Latin dialects. Presented at the Annual meeting of the Linguistic Society of America, Atlanta.
Nevins, Andrew and Bert Vaux. (2004a) Consonant harmony in Karaim, in *Proceedings of the Workshop on Altaic Formal Linguistics*, ed. by A. Csirmaz, Y. Lee, and M. Walter, MIT Working Papers in Linguistics, 46, 175–194, Cambridge, MA.
Nevins, Andrew and Bert Vaux. (2004b) The transparency of contrastive segments in Sibe and Sanjiazi Manchu: Evidence for relativized locality. Presented at GLOW 27, Thessaloniki, April 2004.
Nevins, Andrew and Bert Vaux. (2007) Metalinguistic, shmetalinguistic: The phonology of *shm*-reduplication, in *The Main Session: Papers from the 39th Annual Meeting of the Chicago Linguistic Society*, ed. by Jon E. Cihlar, Amy L. Franklin, David W. Kaiser, and Irene Kimbara, 702–721, Chicago, Chicago Linguistic Society.
Nevins, Andrew and Bert Vaux. (2008) Underlying representations that do not minimize grammatical violations, in *Freedom of Analysis*, ed. by Sylvia Blaho, Patrik Bye, and Martin Krämer, 35–61, Berlin: Mouton de Gruyter.
Newman, John and Robert G. Petterson. (1990) The tones of Kairi. *Oceanic Linguistics* 291, 49–76.
Newman, Paul. (1968) The reality of morphophonemes. *Language* 44, 507–515.
Newman, Paul. (1972) Syllable weight as a phonological variable. *Studies in African Linguistics* 3, 301–323.
Newman, Paul. (1986) Tone and affixation in Hausa. *Studies in African Linguistics* 17, 249–267.
Newman, Paul. (1992) The drift from the coda into the syllable nucleus in Hausa. *Diachronica* 9, 227–238.

Newman, Paul. (2000) *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven, CT: Yale University Press.

Newman, Rochelle S., James R. Sawusch, and Paul A. Luce. (1999) Underspecification and phoneme frequency in speech perception, in *Papers in Laboratory Phonology V: Acquistion and the Lexicon*, ed. by Michael B. Broe and Janet B. Pierrchumbert, 298–311, Cambridge: Cambridge University Press.

Newman, Stanley. (1944) *Yokuts Language of California*. New York: Viking Fund Publications.

Newman, Stanley. (1947) Bella Coola I: Phonology. *International Journal of American Linguistics* 13, 129–134.

Newmeyer, Frederick. (2003) Grammar is grammar and usage is usage. *Language* 79, 682–707.

Newmeyer, Frederick. (2005) *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Newmeyer, Frederick. (2007) Linguistic typology requires crosslinguistic formal categories. *Linguistic Typology* 11, 133–157.

Newport, Elissa L. and Richard N. Aslin. (2004) Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48, 127–162.

Newton, B. E. (1971) Ordering paradoxes in phonology. *Journal of Linguistics* 7, 31–53.

Ní Chiosáin, Máire. (1990) Syllabification, compensatory lengthening and epenthesis in Irish, in *University of Massachusetts Occasional Papers 14*, 67–90.

Ní Chiosáin, Máire and Jaye Padgett. (1997) Markedness, segment realization, and locality in spreading. Technical Report LRC-97-01, Linguistics Research Center, University of California, Santa Cruz.

Ní Chiosáin, Máire and Jaye Padgett. (2001) Markedness, segment realization, and locality in spreading, in *Segmental Phonology in Optimality Theory*, ed. by Linda Lombardi, Cambridge: Cambridge University Press.

Ní Chiosáin, Máire and Jaye Padgett. (2009) Contrast, comparison sets, and the perceptual space, in *Phonological Argumentation: Essays on Evidence and Motivation*, ed. by Steve Parker, London: Equinox Publishing.

Niepokuj, Mary Katherine. (1997) *The Development of Verbal Reduplication in Indo-European*. Washington, DC: Institute for the Study of Man.

Nikiema, Emmanuel and Parth Bhatt. (2005) R diphthongs in French lexifier creoles: Evidence from Haitian. *Toronto Working Papers in Linguistics* 24, 45–63.

Nimmo, Lisa M. and Steven Roodenrys. (2002) Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics* 23, 643–659.

Nishimura, Kohei. (2003) Lyman's Law in Loanwords. Tokyo: Tokyo University.

Niyogi, Partha and Robert Berwick. (1996) A language learning model for finite parameter spaces, in *Computational Approaches to Language Acquisition*, ed. by Michael R. Brent, 605–622, Cambridge, MA: MIT Press.

Nolan, Francis. (1992) The descriptive role of segments: Evidence from assimilation, in *Papers in Laboratory Phonology II*, ed. by Gerard J. Docherty and D. Robert Ladd, 261–280, Cambridge: Cambridge University Press.

Noonan, Michael. (1992) *A Grammar of Lango*. Berlin and New York: Mouton de Gruyter.

Nooteboom, Sieb. (1997) The prosody of speech: Melody and rhythm, in *The Handbook of Phonetic Sciences*, ed. by William J. Hardcastle and John Laver, 640–673, Oxford: Blackwell.

Norman, Kenneth A. and Randall C. O'Reilly. (2003) Modeling hippocampal and neocortical contributions to

recognition memory: A complementary learning systems approach. *Psychological Review* 110, 611–646.
Norris, Dennis, James M. McQueen, and Anne Cutler. (2000) Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23, 299–370.
Nosofsky, Robert M. (1986) Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115, 39–57.
Nosofsky, Robert M. (1990) Relations between exemplar similarity and likelihood models of classification. *Journal of Mathematical Psychology* 34, 393–418.
Noyer, Rolf. (1994) Mobile affixes in Huave: Optimality and morphological wellformedness, in *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, ed. by E. Duncan, D. Farkas, and P. Spaelti, 67–82, Stanford: Stanford Linguistics Association.
Noyer, Rolf. (2007) Spanish loanwords in Huave. Presented at Laval University, Quebec City, Canada.
Nunberg, Geoffrey. (1980) A falsely reported merger in eighteenth-century English: A study in diachronic variation, in *Locating Language in Time and Space*, ed. by William Labov, 221–250, New York: Academic Press.
Odden, David. (1977) Overspecification and variables in phonology. *Linguistic Analysis* 3, 177–196.
Odden, David. (1980) The irrelevancy of the relevancy condition: Evidence for the feature specification constraint. *Linguistic Analysis* 6, 261–394.
Odden, David. (1981) *Problems in Tone Assignment in Shona*. Urbana, IL: University of Illinois dissertation.
Odden, David. (1982) Tonal phenomena in Kishambaa. *Studies in African Linguistics* 13, 177–208.
Odden, David. (1986) On the role of the Obligatory Contour Principle in phonological theory. *Language* 62, 353–383.
Odden, David. (1987) Kimatuumbi phrasal phonology. *Phonology Yearbook* 4, 13–36.
Odden, David. (1988a) Anti antigemination and the OCP. *Linguistic Inquiry* 19, 451–475.
Odden, David. (1988b) Predictable tone systems in Bantu, in *Autosegmental Studies on Pitch Accent*, ed. by Harry van der Hulst and Norval Smith, 225–251, Dordrecht: Foris.
Odden, David. (1990) Syntax, lexical rules and postlexical rules in Kimatuumbi, in *The Phonology-Syntax Connection*, ed. by Sharon Inkelas and Draga Zec, 259–277, Chicago: The University of Chicago Press.
Odden, David. (1991) Vowel geometry. *Phonology* 8, 261–289.
Odden, David. (1993) Interaction between modules in lexical phonology, in *Studies in Lexical Phonology*, ed. by Sharon Hargus and Ellen Kaisse, 111–144, Academic Press.
Odden, David. (1994a) Adjacency parameters in phonology. *Language* 70, 289–330.
Odden, David. (1994b) Syntactic and semantic conditions in KiKongo phrasal phonology, in *Perspectives in Phonology*, ed. by Jennifer Cole and Charles Kisseberth, 167–202, Stanford, CA: CSLI.
Odden, David. (1996) *The Phonology and Morphology of Kimatuumbi*. Oxford: Clarendon Press.
Odden, David. (1999) Typological issues in tone and stress in Bantu, in *Cross-linguistic Studies of Tonal Phenomena, Tonogenesis, Typology and Related Topics*, ed. by Shigeki Kaji, 187–215, Tokyo University of Foreign Studies: Institute for the Study of Languages and Cultures.
Odden, David. (2000) The phrasal tonology of Zinza. *Journal of African Languages and Linguistics* 21, 45–75.

Odden, David. (2006) Minimality and onsetless syllables in Zinza. *Phonology* 23, 431–441.
Ohala, Diane K. (1996) *Cluster Reduction and Constraints in Acquisition*. University of Arizona dissertation.
Ohala, Diane K. (1999) The influence of sonority on children's cluster reductions. *Journal of Communication Disorders* 32, 397–422.
Ohala, John J. (1970) *Aspects of the Control and Production of Speech*. Working Papers in Phonetics, No. 15, Department of Linguistics, UCLA.
Ohala, John J. (1975) Phonetic explanations for nasal sound patterns, in *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, ed. by Charles A. Ferguson, Larry M. Hyman, and John J. Ohala, 289–316, Stanford, CA: Language Universals Project, Department of Linguistics, Stanford University.
Ohala, John J. (1981) The listener as the source of sound change, in *Papers from the Parasession on Language and Behavior*, ed. by Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller, 178–203, Chicago, IL: Chicago Linguistic Society.
Ohala, John J. (1983a) Cross-language use of pitch: An ethological view. *Phonetica* 40, 1–18.
Ohala, John J. (1983b) The origin of sound patterns in vocal tract constraints, in *The Production of Speech*, ed. by Peter F. MacNeilage, 189–216, New York: Springer.
Ohala, John J. (1986) Consumer's guide to evidence in phonology. *Phonology Yearbook* 3, 3–26.
Ohala, John J. (1989) Sound change is drawn from a pool of synchronic variation, in *Language Change: Contributions to the Study of its Causes*, ed. by Leiv Egil Breivik and Ernst Håkon Jahr, 173–198, Berlin: Mouton de Gruyter.
Ohala, John J. (1990a) The phonetics and phonology of assimilation, in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, ed. by John Kingston and Mary E. Beckman, 258–275, Cambridge: Cambridge University Press.
Ohala, John J. (1990b) There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* 18, 153–171.
Ohala, John J. (1992) The segment: Primitive or derived?, in *Papers in Laboratory Phonology II*, ed. by Gerard J. Docherty and D. Robert Ladd, 166–189, Cambridge: Cambridge University Press.
Ohala, John J. (1993) The phonetics of sound change, in *Historical Linguistics: Problems and Perspectives*, ed. by Charles Jones, 237–278, London: Longman.
Ohala, John J. (1994a) Hierarchies of environments for sound variation; plus implications for "neutral" vowels in vowel harmony. *Acta Linguistica Hafniensia* 27, 371–382.
Ohala, John J. (1994b) Towards a universal, phonetically-based theory of vowel harmony, in *Proceedings of ICSLP 94 (International Conference on Spoken Language Processing)*, 491–494, Yokohama, Japan.
Ohala, John J. (2005) Phonetic explanations for sound patterns: Implications for grammars of competence, in *A Figure of Speech: A Festschrift for John Laver*, ed. by W. J. Hardcastle and J. M. Beck, 23–38, London: Erlbaum.
Ohala, John J. and Manjari Ohala. (1986) Testing hypotheses regarding the psychological manifestation of morpheme structure constraints, in *Experimental Phonology*, ed. by John J. Ohala and Jeri J. Jaeger, 239–252, San Diego, CA: Academic Press.
Ohala, Manjari. (1983c) *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.
Öhman, Sven E. G. (1967) Word and sentence intonation: A quantitative model. Technical Report 2-3, Royal

Institute of Technology, Speech Transmission Laboratory Quarterly Progress and Status Report. Pages, 20–54.
O'Keefe, Michael. (2003) Akan vowel harmony. Undergraduate thesis, Swarthmore College.
Okrent, Arika. (2002) A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics, in *Modality in Language and Linguistic Theory*, ed. by Richard Meier, D. Quinto-Pozos, and K. Cormier, 167–174, Cambridge: Cambridge University Press.
Olawsky, Knut J. (2006) *A Grammar of Urarina*. Berlin: Mouton de Gruyter.
Oller, D. K., L. A. Wiemanand, W. J. Doyle, and C. Ross. (1976) Infant babbling and speech. *Journal of Child Language* 3, 1–11.
Oncina, Jose, Pedro Garcia, and Enrique Vidal. (1993) Learning subsequential transducers for pattern recognition interpretation tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, 448–458.
Onishi, Kristine H., Kyle E. Chambers, and Cynthia Fisher. (2002) Learning phonotactic constraints from brief auditory experience. *Cognition* 83, B13–B23.
Oostdijk, Nilleke. (2000) The Spoken Dutch Corpus Project. *The ELRA Newsletter* 5, 4–8.
Oostendorp, Marc van. (1995) *Vowel Quality and Phonological Projection*. Tilburg Dissertation in Language Studies.
Oostendorp, Marc van. (1997) Style levels in conflict resolution, in *Variation, Change, and Phonological Theory*, ed. by Frans Hinskens, Roeland van Hout, and Leo Wetzels, 207–229, Amsterdam: Benjamins.
Oostendorp, Marc van. (2008) Incomplete devoicing in form phonology. *Lingua* 118, 1362–1374.
O'Reilly, Randall C. and Jerry W. Rudy. (2001) Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review* 108, 311–345.
Orešnik, Janez. (1978) The age and importance of the Modern Icelandic type klifr, in *The Nordic Language and Modern Linguistics*, ed. by J. Weinstock, 468–471, Austin, TX: University of Texas Press.
Orgun, Cemil Orhan. (1996) *Sign-based Morphology and Phonology: With Special Attention to Optimality Theory*. University of California, Berkeley dissertation.
Orgun, Cemil Orhan and Sharon Inkelas. (2002) Reconsidering bracket erasure, in *Yearbook of Morphology 2001*, ed. by Geert Booij and Jaap van Marle, 115–146, Dordrecht: Kluwer.
Orgun, Cemil Orhan and Ronald Sprouse. (1999) From MParse to control: Deriving ungrammaticality. *Phonology* 16, 191–224.
Osborn, Henry. (1966) Warao I: Phonology and morphophonemics. *International Journal of American Linguistics* 32, 108–123.
Ostendorf, Mari, Patti J. Price and Stafanie Shattuck-Hufnagel. (1995) The Boston University Radio News Corpus. *Boston University Technical Report, ECS-95-001*, University of Boston.
Ota, Mitsuhiko. (1999) *Phonological Theory and the Acquisition of Prosodic Structure: Evidence from Child Japanese*. Georgetown University dissertation.
Ota, Mitsuhiko. (2006) Input frequency and word truncation in child Japanese: Structural and lexical effects. *Language and Speech* 49, 261–295.
Padden, Carol. (1983) *Interaction of Morphology and Syntax in American Sign Language*. University of California, San Diego dissertation.
Padden, Carole, I. Meir, M. Aronoff, and W. Sandler. (2010) The grammar of space in two new sign languages, in *Sign Languages: A Cambridge Language Survey*, ed. by Diane Brentari, 573–595, Cambridge: Cambridge University Press.

Padgett, Jaye. (1994) Stricture and nasal place assimilation. *Natural Language and Linguistic Theory* 12, 465–513.
Padgett, Jaye. (1995a) Feature classes, in *University of Massachusetts Occasional Papers: Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh-Dickey, and Suzanne Urbanczyk, volume 18, 385–420, Amherst, MA: GLSA.
Padgett, Jaye. (1995b) Partial class behavior and nasal place assimilation, in *Proceedings of the South Western Optimality Theory Workshop*, ed. by Keiichiro Suzuki and Dirk Elzinga, 145–183, University of Arizona Coyote Papers.
Padgett, Jaye. (2002) Feature classes in phonology. *Language* 78, 81–110.
Padgett, Jaye. (2003) Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21, 39–87.
Padgett, Jaye. (2008) Glides, vowels, and features. *Lingua* 118, 1937–1955.
Pak, Marjorie. (2008) *The Postsyntactic Derivation and its Phonological Reflexes*. University of Pennsylvania dissertation.
Pak, Marjorie and Michael Friesner. (2006) French phrasal phonology in a derivational model of PF, in *Proceedings of the 36th Annual Meeting of the North East Linguistic Society*, ed. by Christopher Davis *et al.*, volume 2, 430–49, Amherst, MA: GLSA Publications.
Palmer, Frank R. (1956) "Openness" in Tigre: A problem in prosodic statement. *Bulletin of the School of Oriental and African Studies* 18, 561–577.
Palmer, Stephen E. and Ruth Kimchi. (1986) The information processing approach to cognition, in *Approaches to Cognition: Contrasts and Controversies*, ed. by Terry J. Knapp and Lynn C. Roberts, 37–77, Hillsdale, NJ: Lawrence Erlbaum.
Paolillo, John C. (2002) *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CLSI.
Papoušek, Hanus, Mechtild Papoušek, and David Symmes. (1991) The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development* 14, 415–440.
Paradis, Carole. (1986) *Phonologie et morphologie lexicales: les classes nominales en peul (fula)*. Université de Montréal dissertation.
Paradis, Carole. (1988a) On constraints and repair strategies. *The Linguistic Review* 6, 71–97.
Paradis, Carole. (1988b) Towards a theory of constraint violations. *McGill Working Papers in Linguistics* 5, 1–43.
Paradis, Carole. (1992) *Lexical Phonology and Morphology: The Nominal Classes in Fula*. Garland Publishing.
Paradis, Carole. (2006) The unnatural /Cju/ (< foreign /Cy/) sequence in Russian loanwords: A problem for the perceptual view. *Lingua* 116, 976–995.
Paradis, Carole and Renée Béland. (2002) Syllabic constraints and constraint conflicts in loanword adaptations, aphasic speech and children's errors, in *Phonetics, Phonology and Cognition*, ed. by Jacques Durand and Bernard Laks, 191–225, Oxford: Oxford University Press.
Paradis, Carole and Darlene LaCharité. (1997) Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33, 379–430.
Paradis, Carole and Darlene LaCharité. (2001) Guttural deletion in loanwords. *Phonology* 18, 255–300.
Paradis, Carole and Darlene LaCharité. (2002) Delimiting phonetic approximation in loanword adaptation. Presented at North American Phonology Conference, (NAPhC), Concordia University, Montreal.
Paradis, Carole and Darlene LaCharité. (2006) French loanwords in Arabic, in *Encyclopedia of Arabic Linguistics*, ed. by Kees Versteegh, volume 2, Leiden: Brill Academic Publishers.
Paradis, Carole and Darlene LaCharité. (2008) Apparent phonetic approximation: English loanwords in

Old Quebec French. *Journal of Linguistics* 44, 87–128.
Paradis, Carole and Darlene LaCharité. (2011) Structure Preservation: The resilience of distinctive information, in *Companion to Phonology*, ed. by Marc van Oostendorp *et al.*, 1787–1810, Oxford: Wiley-Blackwell.
Paradis, Carole, Darlene LaCharité, and Frédérick Brault. (1999) L'élision extraordinaire des gutturales, pharyngales et laryngales dans les emprunts et le principe de la non-disponibilité. *Canadian Journal of Linguistics* 44, 149–193.
Paradis, Carole and Caroline Lebel. (2004) Contrasts from segmental parameter settings in loanwords: core and periphery in Quebec French. *Toronto Working Papers in Linguistics* 13, 75–95.
Paradis, Carole, Caroline Lebel, and Darlene LaCharité. (1993) Adaptation d'emprunts: les conditions de la préservation segmentale, in *Proceedings of the Annual Conference of the Canadian Linguistics Association*, ed. by Carrie Dyck, 461–476, Toronto Working Papers in Linguistics.
Paradis, Carole and Jean-François Prunet. (1989) On coronal transparency. *Phonology* 6, 317–348.
Paradis, Carole and François Prunet (eds.) (1991) *The Special Status of Coronals: Internal and External Evidence*. San Diego, CA: Academic Press.
Paradis, Carole and Jean-François Prunet. (2000) Nasal vowels as two segments: Evidence from borrowings. *Language* 76, 324–357.
Paradis, Carole and Mélanie Thibeault. (2004) L'adaptation de la voyelle /y/ en /ju/ dans les emprunts en russe: un cas de palatalisation, non de biphonémicité. *Canadian Journal of Linguistics* 49, 175–222.
Paradis, Carole and Antoine Tremblay. (2009) Nondistinctive features in loanword adaptation: The unimportance of English aspiration in Mandarin Chinese phoneme categorization, in *Studies in Loan Phonology*, ed. by Andrea Calabrese and W. Leo Wetzels, 211–224, Amsterdam: Benjamins.
Parker, Steven. (1997) An OT account of laryngealization in Cuzco Quechua. *Working Papers of the Summer Institute of Linguistics* 41, 1–11.
Parker, Steven. (1999) On the behavior of definite articles in Chamicuro. *Language* 3, 552–562.
Parker, Steven. (2002) *Quantifying the Sonority Hierarchy*. University of Massachusetts dissertation.
Parkinson, Frederick. (1996) *The Representation of Vowel Height in Phonology*. Ohio State University dissertation.
Paster, Mary. (2004) Vowel height harmony and blocking in Buchan Scots. *Phonology* 21, 359–407.
Paster, Mary. (2005) Pulaar verbal extensions and phonologically driven affix order, in *Yearbook of Morphology*, ed. by Geert Booij and Jaap van Marle, 155–199, Springer.
Paster, Mary. (2006) *Phonological Conditions on Affixation*. University of California, Berkeley dissertation.
Paster, Mary and Yuni Kim. (2007) Downstep in Tiriki. Manuscript Pomona College and University of California, Berkeley.
Pater, Joe. (1997) Minimal violation and phonological development. *Language Acquisition* 6, 201–253.
Pater, Joe. (2000) Nonuniformity in English Stress: The role of ranked and lexically specific constraints. *Phonology* 17, 237–274.
Pater, Joe. (2004) Bridging the gap between perception and production with minimally violable constraints, in *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 219–244, Cambridge: Cambridge University Press.
Pater, Joe. (2008a) Gradual learning and convergence. *Linguistic Inquiry* 39, 334–345.

Pater, Joe. (2008b) Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17, 237–274.

Pater, Joe. (2009a) Morpheme-specific phonology: constraint indexation and inconsistency resolution, in *Phonological Argumentation: Essays on Evidence and Motivation*, ed. by Steven Parker, London: Equinox.

Pater, Joe. (2009b) Review of "*The Harmonic Mind*," by Smolensky and Legendre. *Phonology* 26, 217–226.

Pater, Joe. (2009c) Weighted constraints in generative linguistics. *Cognitive Science* 33, 999–1035.

Pater, Joe and Jessica A. Barlow. (2003) Constraint conflict in cluster reduction. *Journal of Child Language* 30, 487–526.

Pater, Joe, Rajesh Bhatt, and Christopher Potts. (2007) Linguistic optimization. http://people.umass.edu/pater/pater-bhatt-potts-hg07.pdf.

Pater, Joe, Christopher Potts, and Rajesh Bhatt. (2006) Harmonic grammar with linear programming. Rutgers Optimality Archive 872.

Pater, Joe, Christine Stager, and Janet Werker. (2004) The perceptual acquisition of phonological contrasts. *Language* 80, 384–402.

Pater, Joe and Anne-Marie Tessier. (2003) Phonotactic knowledge and the acquisition of alternations, in *Proceedings of the 15th International Congress on Phonetic Sciences, Barcelona*, ed. by M. J. Solé, D. Recasens, and J. Romero, 1777–1180.

Pater, Joe and Anne-Michelle Tessier. (2006) L1 Phonotactic knowledge and the L2 acquisition of alternations, in *Inquiries in Linguistic Development: Studies in Honor of Lydia White*, ed. by R. Slabakova, S. Montrul, and P. Prevost, 115–131, Amsterdam: Benjamins.

Pater, Joe and Adam Werle. (2001) Typology and variation in child consonant harmony, in *Proceedings of HILP5*, ed. by Caroline Féry, Antony Dubach Green, and Ruben van de Vijver, 119–139, Potsdam, University of Potsdam.

Pater, Joe and Adam Werle. (2003) Direction of assimilation in child consonant harmony. *Canadian Journal of Linguistics* 48, 385–408.

Patrick, Peter L. (1992) Creoles at the intersection of variable processes: *-t,-d* deletion and past-marking in the Jamaican mesolect. *Language Variation and Change* 3, 171–189.

Patterson, David and Cynthia M. Connine. (2001) Variant frequency in flap production. *Phonetica* 58, 254–274.

Paul, Hermann. (1880) *Prinzipien der Sprachgeschichte*. Halle: Niemeyer. English translation of 2nd (1886) edition, *Principles of the History of Language*, published by McGroth Publishing Company, College Park, 1970.

Paulian, Christiane. (1975) *Le kukuya: langue teke du Congo*. Paris: SELAF.

Pearce, Mary. (2005) Kera tone and voicing, in *UCL Working Papers in Linguistics*, 17, 61–82.

Pearce, Mary. (2007a) Iambicity in Kera, in *Selected Proceedings of the 37th Annual Conference on African Linguistics*, ed. by D. Payne and J. Peña, 66–76, Somerville, MA: Cascadilla Press.

Pearce, Mary. (2007b) *The Interaction of Tone with Voicing and Foot Structure: Evidence from Kera Phonetics and Phonology*. University of London dissertation.

Peirce, Charles Sanders. (1958) *Collected Papers of Charles Sander Peirce, 1931–1935*. Cambridge, MA: Harvard University Press.

Peperkamp, Sharon. (1996) On the prosodic representation of clitics, in *Interfaces in Phonology*, ed. by Ursula Kleinhenz, 104–128, Berlin: Akademie Verlag.

Peperkamp, Sharon. (1997) *Prosodic Words*. HIL Dissertation Series 34, The Hague: Holland Academic Graphics.

Peperkamp, Sharon. (2007) Do we have innate knowledge about phonological markedness? Comments on Berent, Steriade, Lennertz, and Vaknin. *Cognition* 104, 631–637.

Peperkamp, Sharon, Rozenn Le Calvez, Jean-Pierre Nadal, and Emmanuel Dupoux. (2006a) The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101, B31–B41.

Peperkamp, Sharon and Emmanuel Dupoux. (2002) Loanword adaptations: Three problems for phonology and a psycholinguistic solution. Presented at the North American Phonology Conference (NAPhC), Concordia University, Montreal.

Peperkamp, Sharon and Emmanuel Dupoux. (2003) Reinterpreting loanword adaptations: The role of perception, in *Proceedings of the 15th International Congress of the 15th International Congress of Phonetic Sciences*, 367–370.

Peperkamp, Sharon and Emmanuel Dupoux. (2007) Learning the mapping from surface to underlying representations in an artificial language, in *Laboratory Phonology IX*, ed. by Jennifer Cole and José Hualde, 315–338, Berlin: Mouton de Gruyter.

Peperkamp, Sharon, Katrin Skoruppa, and Emmanuel Dupoux. (2006b) The role of phonetic naturalness in phonological acquisition, in *Proceedings of the 30th Annual Boston University Conference on Language Development*, ed. by D. Bamman, T. Magnitskaia, and C. Zaller, 464–475, Sommerville, MA: Cascadilla Press.

Percival, Keith. (1976) On the historical source of immediate-constituent analysis, in *Notes from the Linguistic Underground*, ed. by James McCawley, 229–242, New York: Academic Press.

Pereira, Fernando. (2000) Formal grammar and information theory: together again? *Philosophical Transactions of the Royal Society of London A*, 358, 1239–1253.

Perlmutter, David. (1971) *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart, and Winston.

Perlmutter, David. (1992) Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23, 407–442.

Perruchet, Pierre and Ronald Peereman. (2004) The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics* 17, 97–119.

Pertz, Doris L. and Thomas G. Bever. (1975) Sensitivity to phonological universals in children and adolescents. *Language* 51, 149–162.

Peters, P. Stanley and R. W. Ritchie. (1973) On the generative power of transformational grammars. *Information Sciences* 6, 49–83.

Peterson, Gordon and Ilse Lehiste. (1960) Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693–703.

Petitto, Laura A. (2000) On the biological foundations of human language, in *The Signs of Language Revisited: An Anthology in Honor of Ursula Bellugi and Edward Klima*, ed. by K. Emmorey and H. Lane, 447–471, Mahwah, NJ: LEA.

Petitto, Laura A. and Paula Marantette. (1991) Babbling in the manual mode: Evidence for the ontogeny of language. *Science* 251, 1493–1496.

Philippson, Gérard. (1991) *Tons et accent dans les langues Bantu d'Afrique orientale: étude comparative typologique et diachronique*. Paris: Universite René Descartes dissertation.

Philippson, Gérard. (1998) Tone reduction vs. metrical attraction in the evolution of Eastern Bantu tone systems, in *Theoretical Aspects of Bantu Tone*, ed. by Larry M. Hyman and Charles W. Kisseberth, 315–329, Stanford, CA: CSLI.

Phillips, Betty S. (1984) Word frequency and the actuation of sound change. *Language* 60, 320–342.

Phillips, Betty S. (2001) Lexical diffusion, lexical frequency, and lexical analysis, in *Frequency and the Emergence of Linguistic*

Structure, ed. by Joan L. Bybee and Paul Hopper, 123–136, Benjamins.
Phillips, Betty S. (2006) *Word Frequency and Lexical Diffusion*. New York: Palgrave Macmillan.
Picard, Marc. (1984) English flapping and aspiration revisted. *Canadian Journal of Linguistics* 29, 42–57.
Pickering, M. J. and S. Garrod. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27, 169–226.
Pierrehumbert, Janet B. (1980) *The Phonology and Phonetics of English Intonation*. MIT dissertation.
Pierrehumbert, Janet B. (1990) Phonological and phonetic representation. *Journal of Phonetics* 18, 375–394.
Pierrehumbert, Janet B. (1994) Syllable structure and word structure: A study of triconsonantal clusters in English, in *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, ed. by Patricia A. Keating, 168–188, Cambridge: Cambridge University Press.
Pierrehumbert, Janet B. (1999) What people know about sounds of language. *Studies in the Linguistic Sciences* 29, 111–120.
Pierrehumbert, Janet B. (2000) The phonetic grounding of phonology. *Bulletin de la communication parlée* 5, 7–23.
Pierrehumbert, Janet B. (2001a) Exemplar dynamics: Word frequency, lenition and contrast, in *Frequency Effects and the Emergence of Linguistic Structure*, ed. by Joan L. Bybee and Paul Hopper, 137–157, Amsterdam: Benjamins.
Pierrehumbert, Janet B. (2001b) Why phonological constraints are so coarse-grained. *Language and Cognitive Processes* 16, 691–698.
Pierrehumbert, Janet B. (2002) Word-specific phonetics, in *Papers in Laboratory Phonology VII*, ed. by Carlos Gussenhoven and Natasha Warner, 101–139, Berlin: Mouton de Gruyter.
Pierrehumbert, Janet B. (2003) Probabilistic phonology: Discrimination and robustness, in *Probabilistic Linguistics*, ed. by Rens Bod, Jennifer Hay, and S. Jannedy, 177–228, Cambridge, MA: MIT Press.
Pierrehumbert, Janet B. (2006a) The next toolkit. *Journal of Phonetics* 34, 516–530.
Pierrehumbert, Janet B. (2006b) The statistical basis of an unnatural alternation, in *Laboratory Phonology VIII: Varieties of Phonological Competence*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 81–107, Berlin: Mouton de Gruyter.
Pierrehumbert, Janet B. and Mary E. Beckman. (1988) *Japanese Tone Structure*. Cambridge, MA: MIT Press.
Pierrehumbert, Janet B., Mary E. Beckman, and D. Robert Ladd. (1996) Laboratory phonology, in *Current Trends in Phonology: Models and Methods. Volume 2*, ed. by J. Durand and B. Laks, 535–548, Salford, Manchester: European Studies Research Institute.
Pierrehumbert, Janet B., Mary E. Beckman, and D. Robert Ladd. (2000) Conceptual foundations of phonology as a laboratory science, in *Phonological knowledge: Conceptual and Empirical Issues*, ed. by N. Burton-Roberts, P. Carr, and G. J. Docherty, 273–304, Oxford: Oxford University Press.
Pierrehumbert, Janet B. and Julia Hirschberg. (1990) The meaning of intonation contours in the interpretation of discourse, in *Intentions in Communication*, ed. by Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, 271–311, Cambridge, MA: MIT Press.
Pierrehumbert, Janet B. and Rami Nair. (1995) Word games and syllable structure. *Language and Speech* 38, 77–114.
Piggott, Glyne. (1980) *Aspects of Odawa Morphophonemics*. New York: Garland Publishing.
Piggott, Glyne. (1988a) The parameters of nasalization. *McGill Working Papers in Linguistics* 5, 128–177.

Piggott, Glyne. (1988b) A parametric approach to nasal harmony, in *Features, Segmental Structure, and Harmony Processes*, ed. by Harry van der Hulst and Norval Smith, 131–167, Dordrecht: Foris.

Piggott, Glyne. (1992) Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* 10, 33–78.

Piggott, Glyne. (1996) Implications of consonant nasalization for a theory of harmony. *Canadian Journal of Linguistics* 41, 141–174.

Piggott, Glyne. (2003) Theoretical implications of segment neutrality in nasal harmony. *Phonology* 20, 375–424.

Piggott, Glyne and Harry van der Hulst. (1997) Locality and the nature of nasal harmony. *Lingua* 103, 85–112.

Piggott, Glyne and Rajendra Singh. (1985) The phonology of epenthetic segments. *Canadian Journal of Linguistics* 30, 415–451.

Pike, Eunice V. and Priscilla Small. (1974) Downstepping terrace tone in Coatzospan Mixtec, in *Advances in Tagmemics*, ed. by Ruth Brend, 105–134, Amsterdam: North-Holland Publishing Company.

Pike, Eunice V. and Kent Wistrand. (1974) Step-up terrace tone in Acatlán Mixtec, in *Advances in Tagmemics*, ed. by Ruth Brend, 83–104, Amsterdam: North-Holland Publishing Company.

Pike, Kenneth L. (1943a) *Phonetics*. Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. (1943b) Taxemes and immediate constituents. *Language* 19, 65–82.

Pike, Kenneth L. (1945) *The Intonation of American English.* Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. (1947) *Phonemics: A Technique for Reducing Languages to Writing*. Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. (1948) *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language.* Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. (1967) *Language in Relation to a Unified Theory of the Structure of Human Behaviour*. The Hague: Mouton.

Pike, Kenneth L. and Eunice V. Pike. (1947) Immediate constituents of Mazatec syllables. *IJAL* 13, 78–91.

Piñeros, Carlos. (1998) *Prosodic Morphology in Spanish: Constraint Interaction in Word Formation*. Ohio State University dissertation.

Pinker, Steven. (1991) Rules of language. *Science* 153, 530–535.

Pinker, Steven and Alan Prince. (1988) On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193.

Pitt, M. A., K. Johnson, Elizabeth V. Hume, S. Kiesling, and W. Raymond. (2005) The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability. *Speech Communication* 45, 89–95.

Pitt, Mark A. and James M. McQueen. (1998) Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39, 347–370.

Pitt, Mark A. and Arthur G. Samuel. (1995) Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology* 29, 149–188.

Pittman, Richard S. (1948) Nuclear structure in linguistics. *Language* 24(3), 287–292.

Plag, Ingo, Gero Kunter, and Sabine Lappe. (2007) Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics and Linguistic Theory* 3, 199–232.

Plag, Ingo, Gero Kunter, Sabine Lappe, and Maria Braun. (2008) The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language* 84, 760–794.

Plauché, Madelaine, Cristina Delogu, and John J. Ohala. (1997) Asymmetries in consonant confusion, in *Proceedings of EUROSPEECH '97*, volume 4, 2187–2190.

Plaut, David C. and Christopher T. Kello. (1999) The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach, in *The Emergence of Language*, ed. by Brian MacWhinney, 381–415, Mahwah, NJ: Lawrence Erlbaum Associates.

Plénat, Marc. (1995) Une approche prosodique de la morphologie du verlan. *Lingua* 95, 97–129.

Ploch, Steven. (1999) *Nasals on my Mind: The Phonetic and the Cognitive Approach to the Phonology of Nasality*. School of Oriental and African Studies, University of London dissertation.

Plug, L. (2005) From words to actions: The phonetics of "eigenlijk": in two communicative contexts. *Phonetica* 62, 131–145.

Pluymaekers, M., Mirjam Ernestus, and R. H. Baayen. (2005a) Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62, 146–159.

Pluymaekers, Mark, Mirjam Ernestus, and R. Harald Baayen. (2005b) Frequency and acoustic length: The case of derivational affixes in Dutch. *Journal of the Acoustical Society of America* 118, 2561–2569.

Pöchtrager, Markus Alexander. (2006) *The Structure of Length*. Universität Wien dissertation.

Poizner, Howard, Ursula Bellugi, and Ryan D. Tweney. (1981) Processing of formational, semantic, and iconic information in American Sign Language. *Journal of Experimental Psychology: Human Perception and Performance* 7, 1146–1159.

Polgárdi, Krisztina. (1998) *Vowel Harmony: An Account in Terms of Government and Optimality*. The Hague: Holland Academic Graphics.

Poliquin, Gabriel. (2006) *Canadian French Vowel Harmony.* Harvard University dissertation. ROA 861.

Pollack, Karen, Jackson Gandour, and David Sorensen. (1987) Goose Latin: A study of an idioglossia. *American Speech* 62, 92–94.

Pollard, Carl and Ivan Sag. (1994) *Head-driven Phrase Structure Grammar*, Chicago CSLI and University of Chicago Press.

Poplack, Shana, David Sankoff, and Christopher Miller. (1988) The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26, 47–104.

Poppe, N. N. (1964) *Bakshir Manual: Descriptive Grammar and Texts with a Bashkir-English Glossary*. Bloomington, IN: Indiana University Press.

Port, Robert and Penny Crawford. (1989) Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17, 257–282.

Port, Robert F. and Adam P. Leary. (2005) Against formal phonology. *Language* 81, 927–964.

Port, Robert F. and Michael O'Dell. (1985) Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13, 455–471.

Poser, William J. (1982) Phonological representation and action-at-a-distance, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, 121–158, Dordrecht: Foris.

Poser, William J. (1984) *The Phonetics and Phonology of Tone and Intonation in Japanese*. MIT, dissertation.

Poser, William J. (1992) Blocking out of the lexicon, in *Lexical Matters*, ed. by Ivan Sag and Anna Szabolcsi, 111–130, Stanford, CA: CSLI Publications.

Poser, William J. (1993) Are strict cycle effects derivable?, in *Lexical Phonology and Morphology*, ed. by Ellen Kaisse and Sharon Hargus, 315–321, New York: Academic Press.

Post, Emil. (1943) Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics* 65, 197–215.
Post, Mark William. (2007) *A Grammar of Galo*. La Trobe University dissertation.
Postal, Paul M. (1968) *Aspects of Phonological Theory*. New York: Harper and Row.
Potts, Christopher. (2005) *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
Pouplier, M. (2007) Tongue kinematics during utterances elicited with the SLIP technique. *Language and Speech* 50, 311–341.
Pouplier, M. (2008) The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics* 36, 114–140.
Press, Margaret. (1980) *Chemehuevi: A Grammar and Lexicon*. Berkeley, CA: University of California Press.
Prieto, Pilar. (1997) Prosodic manifestation of syntactic structure in Catalan, in *Issues in the Phonology of the Iberian Languages*, ed. by Fernando Martinez-Gil and Alfonso Morales-Front, 179–199, Washington, DC: Georgetown University Press.
Prieto, Pilar. (2005) Syntactic and eurhythmic constraints on phrasing decisions. *Studia Linguistica* 59, 194–222.
Prieto, Pilar. (2006) The relevance of metrical information in early prosodic word acquisition: A comparison of Catalan and Spanish. *Language and Speech* 49, 231–259.
Prince, Alan. (1975) *The Phonology and Morphology of Tiberian Hebrew*. MIT dissertation.
Prince, Alan. (1983) Relating to the grid. *Linguistic Inquiry* 14, 19–100.
Prince, Alan. (1990) Quantitative consequences of rhythmic organization, in *Chicago Linguistic Society 26: Parasession on the Syllable in Phonetics and Phonology*, 355–398, Chicago: Chicago Linguistic Society.
Prince, Alan. (2006) Implication and impossibility in grammatical systems: What it is and how to find it. Manuscript, Rutgers University. Available on Rutgers Optimality Archive, ROA-880.
Prince, Alan and Paul Smolensky. (1993) Optimality Theory: constraint interaction in generative grammar. ROA 537-0802.
Prince, Alan and Paul Smolensky. (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
Prince, Alan and Bruce Tesar. (2004) Learning phonotactic distributions, in *Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 245, Cambridge: Cambridge University Press.
Prunet, Jean-François, Renée Béland, and Ali Idrissi. (2000) The mental representation of Semitic words. *Linguistic Inquiry* 31, 609–648.
Przezdziecki, Marek. (2005) *Vowel Harmony and Coarticulation in Three Dialects of Yoruba: Phonetics Determining Phonology.* Cornell University dissertation.
Pulgram, Ernst. (1970) *Syllable, Word, Nexus, Cursus*. The Hague: Mouton.
Pulleyblank, Douglas. (1983) *Tone in Lexical Phonology*. MIT dissertation. Published by D. Reidel, Dordrecht, 1986.
Pulleyblank, Douglas. (1988) Underspecification, the feature hierarchy and Tiv vowels. *Phonology* 5, 299–326.
Pulleyblank, Douglas. (1989) Patterns of feature cooccurrence: The case of nasality, in *Proceedings of the Arizona Phonology Conference*, ed. by S. Lee Fulmer, Masahide Ishihara, and Wendy Wiswell, Coyote Papers 9, 98–115.
Pulleyblank, Douglas. (1995) Feature geometry and underspecification, in *Frontiers of Phonology: Atoms, Structures, Derivations*, ed. by Jacques Durand and Francis Katamba, 3–33, Harlow, UK: Longman.
Pulleyblank, Douglas. (1996) Neutral vowels in Optimality Theory: A comparison of Yoruba and Wolof.

Canadian Journal of Linguistics 41, 295–347.

Pulleyblank, Douglas. (2002) Harmony drivers: No disagreement allowed, in *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, 249–267.

Pulleyblank, Douglas. (2009) Patterns of reduplication in Yoruba, in *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, Cambridge, MA: MIT Press.

Pullum, Geoffrey K. (1976) The Duke of York gambit. *Journal of Linguistics* 12, 83–102.

Pulman, Stephen and Mark Hepple. (1993) A feature-based formalism for two-level phonology: A description and implementation. *Computer Speech and Language* 7, 333–358.

Pustejovsky, James. (2000) Events and the semantics of opposition, in *Events as Grammatical Objects: The Converging Perspectives of Lexical Semantics and Syntax*, ed. by C. Tenny and J. Pustejovsky, 445–482, Stanford, CA: Center for the Study of Language and Information.

Pycha, Anne. (2007) Phonetic vs. phonological lengthening in affricates, in *International Congress of Phonetic Sciences XVI*, 1757–1760.

Pycha, Anne. (2009) Lengthened affricates as a test case for the phonetics-phonology interface, in *Journal of the International Phonetic Association 39*, 1–31, timestamp = 2009.12.15.

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted. (2003) Phonological rule learning and its implications for a theory of vowel harmony in *Proceedings of the 22th Annual Meeting of the West Coast Conference on Formal Linguistics*, 101–114, Somerville, MA: Cascadilla Press.

Pycha, Anne, Eurie Shin, and Ryan Shosted. (2006) Directionality of assimilation in consonant clusters: An experimental approach. *University of California Berkeley Phonology Laboratory Annual Report* 152–166.

Pylkkänen, Liina, Andrew Stringfellow, and Alec Marantz. (2002) Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language* 81, 666–678.

Rabiner, Lawrence and Biing-Hwang Juang. (1993) *Fundamentals of Speech Recognition.* Englewood Cliffs, NJ: PTR Prentice-Hall.

Racine, Isabelle and François Grosjean. (2002) La production du E caduc facultatif est-elle prévisible? Un début de réponse. *Journal of French Language Studies* 12, 307–326.

Raimy, Eric. (2000) Remarks on backcopying. *Linguistic Inquiry* 31, 541–552.

Ramstedt, Gustaf John. (1902) *Bergtscheremissische Sprachstudien.* Helsingfors: Druckerei der Finnischen Literaturgesellschaft.

Ranbom, L. J. and C. M. Connine. (2007) Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57, 273–298.

Randolph, Mark A. (1989) *Syllable-based Constraints on Properties of English Sounds*. MIT dissertation.

Rapp, Brenda and Matthew Goldrick. (2000) Discreteness and interactivity in spoken word production. *Psychological Review* 107, 460–499.

Rappaport, Malka. (1984) *Issues in the Phonology of Tiberian Hebrew*. MIT dissertation.

Ratcliff, Roger, Pablo Gomez, and Gail McKoon. (2004) A diffusion model account of the lexical decision task. *Psychological Review* 111, 159–182.

Ratliff, Martha. (1992) *Meaningful Tone: A Study of Tonal Morphology in Compounds, Form Classes and Expressive Phrases in White Hmong*. DeKalb, IL: Northern Illinois University, Center for Southeast Asian Studies.

Rehg, K. and D. Sohl. (1981) *Ponapean Reference Grammar*. Honolulu, HI: University Press of Hawaii.

Reilly, Judy. (2006) How faces come to serve grammar: The development of non-manual morphology in ASL, in *Advances in the Development of Sign Language by Deaf Children*, ed. by B. Schick, M. Marschark, and P. Spencer, 262–290, Oxford: Oxford University Press.

Reinhart, Tanya. (1995) Interface strategies, in *OTS Working Papers in Theoretical Linguistics*, 55–109.

Reinhart, Tanya. (2006) *Interface Strategies: Optimal and Costly Computations*. Cambridge, MA: MIT Press.

Reinholt Peterson, Niels. (1986) Perceptual compensation for segmentally conditioned fundamental-frequency perturbations. *Phonetica* 43, 31–42.

Reiss, Charles. (2003) Quantification in structural descriptions: Attested and unattested patterns. *The Linguistic Review* 20, 305–338.

Rennison, John R. and Friedrich Neubarth. (2003) An X-bar theory of government phonology, in *Living on the Edge: 28 Papers in Honour of Jonathan Kaye*, ed. by Stefan Ploch, 95–130, Berlin: Mouton de Gruyter.

Revithiadou, Anthi and Vassili Spyropoulos. (2005) The multiple spell-out hypothesis and the phonological component: Evidence from Greek, in *Proceedings of the 35th Annual Meeting of the North East Linguistic Society*, ed. by Leah Bateman and Cherlon Ussery, 523–537.

Revithiadou, Anthi and Vassili Spyropoulos. (2009) A dynamic approach to the syntax-phonology interface: A case study from Greek, in *Interphases: Phase-theoretic Investigations of Linguistic Interfaces*, 202–233, Oxford: Oxford University Press.

Reynolds, William. (1994) *Variation and Phonological Theory.* University of Pennsylvania, Philadelphia dissertation.

Rhee, Sang Jik. (2002) *Empty Nuclei in Korean.* University of Leiden dissertation. Published by LOT, No. 61.

Riad, Tomas and Carlos Gussenhoven (eds.) (2007) *Tones and Tunes: Typological Studies in Word and Sentence Prosody*, volume 1, Berlin: Mouton De Gruyter.

Rialland, Annie. (1981) Le système tonal du gurma (langue gur de Haute-Volta). *Journal of African Languages and Linguistics* 3, 39–64.

Rialland, Annie. (1983) Le système tonal du moba comparé à celui du gurma, in *Current Approaches to African Linguistics*, ed. by Jonathan Kaye, Hilda Koopman, Dominque Sportiche, and André Dugas, 217–234, Dordrecht: Foris.

Rialland, Annie. (1986) Schwa et syllabes en français, in *Studies in Compensatory Lengthening*, ed. by Leo Wetzels and Engin Sezer, 187–226, Dordrecht: Foris.

Rialland, Annie. (2001) Anticipatory raising in downstep realization: Evidence for preplanning in tone production, in *Cross-linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*, ed. by Shikegi Kaji, 301–322, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Ribeiro, Amanda. (2000) What's cool for kids: About ZOOM. *Newsday* May 9.

Ribeiro, Eduardo. (2003) Directionality in vowel harmony: The case of Karajá (Macro-Jê), in *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, 475–485.

Rice, Keren. (2000) *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb.* Cambridge: Cambridge University Press.

Rice, Keren. (2005) Liquid Relationships, in *Toronto Working Papers in Linguistics, Volume 24: Special Issue on Similarity in Phonology*, ed. by Chiara Frigeni, Manami Hirayama, and Sara Mackenzie, 31–44.

Richards, Norvin. (2009) *Uttering Trees*. Cambridge, MA: MIT Press.

Ridouane, Rachid. (2008) Syllables without vowels: Phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25, 321–359.
Rietveld, Toni, Roeland van Hout and Mirjam Ernestus. (2004) Pitfalls in corpus research. *Computers and the Humanities* 38, 343–362.
Riggle, Jason. (2003) Nonlocal reduplication, in *Proceedings of the 34th Annual Meeting of the North East Linguistic Society.*
Riggle, Jason. (2004) *Generation, Recognition, and Learning in Finite State Optimality Theory*. UCLA: dissertation.
Riggle, Jason. (2006a) Infixing reduplication in Pima and its theoretical consequences. *Natural Language and Linguistic Theory* 24, 857–891.
Riggle, Jason. (2006b) Using entropy to learn OT grammars from surface forms alone, in *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by D. Baumer, D. Montero, and M. Scanlon, 346–353, Somerville, MA: Cascadilla Press.
Riha, Helena. (2008) *Lettered Words and Roman Letter Characters: A Study of Alphabetic Writing in Chinese Newswires*. Ohio State University dissertation.
Ringen, Catherine O. (1975) *Vowel Harmony: Theoretical Implications*. Indiana University dissertation. Published by Garland Publishing, New York 1988.
Ringen, Catherine O. and Orvakki Heinämäki. (1999) Variation in Finnish vowel harmony: An OT account. *Natural Language and Linguistic Theory* 17, 303–337.
Ringen, Catherine O. and Miklós Kontra. (1989) Hungarian neutral vowels. *Lingua* 78, 181–191.
Ringen, Catherine O. and Robert M. Vago. (1998) Hungarian vowel harmony in Optimality Theory. *Phonology* 15, 393–416.
Ringen, Catherine O. and Robert M. Vago. (2011) Geminates: heavy or long?, in *Handbook of the Syllable*, ed. by Charles Cairns and Eric Raimy, 155–169, Leiden: Brill.
Ritter, Nancy A. (1995) *The Role of Universal Grammar in Phonology: A Government Phonology Approach to Hungarian*. New York University dissertation.
Ritter, Nancy A. (1997) Headedness as a means of encoding stricture, in *Phonology in Progress–Progress in Phonology*, ed. by Geert Booij and J. van der Weijer, 333–365, Leiden: Holland Academic Graphics.
Ritter, Nancy A. (2006a) Georgian consonant clusters: The complexity is in the structure, not the melody. *The Linguistic Review* 23, 429–464.
Ritter, Nancy A. (2006b) Government phonology, in *The Encyclopedia of Language and Linguistics*, ed. by Keith Brown, volume 3, 106–113, Oxford: Elsevier.
Rivard, Adjutor and Louis-Philippe Geoffrion. (1931) *Glossaire du parler français au Canada*. Quebec City: Presses de l'Université Laval.
Rizzi, Luigi. (1997) The fine structure of the left periphery, in *Elements of Grammar: Handbook in Generative Syntax*, ed. by Liliane Haegeman, 75–116, Dordrecht: Kluwer Publishing.
Rizzolo, Olivier. (2007) The syllable is not a valid constituent: Evidence from two Serbo-Croatian language games, in *Formal Approaches to Slavic Linguistics*, 264–281, Ann Arbor, MI: Michigan Slavic Publications.
Roach, P. J. (1982) On the distinction between "stress-timed" and "syllable-timed" languages, in *Linguistic Controversies*, ed. by David Crystal, 73–79, London: Edward Arnold.
Roark, Brian and Katherine Demuth. (2000) Prosodic constraints and the learner's environment: A corpus study, in *Proceedings of the 24th Annual Boston University Conference on Language Development*, ed. by S. C. Howell,

S. A. Fish, and T. Keith-Lucas, 597–608, Somerville, MA: Cascadilla Press.
Roberts, J. and W. Labov. (1995) Learning to talk Philadelphian: Acquisition of short /a/ by preschool children. *Language Variation and Change* 7, 101–112.
Robins, R. H. (1957) Vowel nasality in Sundanese, in *Studies in Linguistic Analysis*, 87–103, Oxford: Blackwell.
Robinson, J. J. (1970) Dependency structures and transformational rules. *Language* 46, 259–285.
Rochemont, Michael. (1986) *Focus in Generative Grammar*. Amsterdam and Philadelphia: Benjamins.
Romani, Cristina and Andrea Calabrese. (1998) Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language* 64, 83–121.
Roodenrys, Steven and Melina Hinton. (2002) Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory and Cognition* 28, 29–33.
Rooth, Mats. (1992) A theory of focus interpretation. *Natural Language Semantics* 1, 75–116.
Rooth, Mats. (1996) Focus, in *The Handbook of Contemporary Semantic Theory*, ed. by Shalom Lappin, 271–297, Oxford: Blackwell.
Rose, Sharon. (1996) Variable laryngeals and vowel-lowering. *Phonology* 13, 73–117.
Rose, Sharon. (2004) Long distance vowel-consonant agreement in Harari. *Journal of African Languages and Linguistics* 25, 41–87.
Rose, Sharon and Lisa King. (2007) Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50, 451–504.
Rose, Sharon and Rachel Walker. (2004) A typology of consonant agreement as correspondence. *Language* 80, 475–531.
Rose, Yvan. (2000) *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology*. McGill University dissertation.
Rose, Yvan, Brian MacWhinney, Rodrigue Byrne, Gregory Hedlund, Keith Maddocks, Philip O'Brien and Todd Wareham. (2006) Introducing Phon: A software solution for the study of phonological acquisition, in *Proceedings of the 30th Boston University Conference on Language Development*, ed. by D. Bamman, T. Magnitskaia, and C. Zaller, 489–500, Somerville, MA: Cascadilla Press.
Rosenbloom, Paul C. (1950) *The Elements of Mathematical Logic*. New York: Dover Publications.
Rosenthall, Sam and Harry van der Hulst. (1999) Weight-by-position by position. *Natural Language and Linguistic Theory* 17, 499–540.
Ross, John Robert. (1967) *Constraints on Variables in Syntax*. MIT dissertation.
Rotenberg, Joel. (1978) *The Syntax of Phonology.* MIT dissertation.
Rousseau, Pascal and David Sankoff. (1978) Advances in variable rule methodology, in *Linguistic Variation: Models and Methods*, ed. by David Sankhoff, 57–69, New York: Academic.
Rowicka, Grazyna. (1999) *On Ghost Vowels: A Strict CV Approach*. The Hague: Holland Academic Graphics.
Rubach, Jerzy. (1990) Final devoicing and cyclic syllabification in German. *Linguistic Inquiry* 21, 79–94.
Rubach, Jerzy. (1996) Shortening and ambisyllabicity in English. *Phonology* 13, 197–237.
Rubach, Jerzy and Geert Booij. (1990a) Edge of constituent effects in Polish. *Natural Language and Linguistic Theory* 8, 427–463.
Rubach, Jerzy and Geert Booij. (1990b) Syllable structure assignment in Polish. *Phonology* 7, 121–58.
Rudin, C. (1980) Phonetic evidence for a phonological rule: G-deletion in Turkish. *Research in Phonetics* 1, 217–232.
Rumelhart, David E. and James L. McClelland. (1987) Learning the past tenses of English verbs: implicit rules

or parallel distributed processing, in *Mechanisms of Language Acquisition*, ed. by Brian MacWhinney, 195–248, Mahwah, NJ: Lawrence Erlbaum Associates.

Rumelhart, David E., James L. McClelland, and the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1: Foundations. Cambridge, MA: Bradford Books/MIT Press.

Rumsey, Alan. (2000) Bunuba, in *Handbook of Australian Languages*, ed. by R. M. W. Dixon and Barry Blake, volume V, 35–152, Oxford: Oxford University Press.

Rupp, James E. (1990) The Lealao Chinantec syllable, in *Syllables, Tone, and Verb Paradigms*, ed. by William R. Merrifield and Calvin R. Rensch, volume 4 of *Studies in Chinantec languages*, 63–73, Dallas, TX: SIL and University of Texas at Arlington.

Russell, Kevin. (1999) MOT: Sketch of an OT approach to morphology. ROA 352.

Russell, Kevin. (2008) Sandhi in Plains Cree. *Journal of Phonetics* 36, 450–464.

Russell, Robert and Dolores Russell. (1959) Syntactotonemics in Amahuaca (Pano). *Série lingüística especial* 1, 128–167.

Russo, Tommaso. (2005) A cross-cultural, cross-linguistic analysis of metaphors in two Italian Sign Language Registers. *Sign Language Studies* 5, 333–359.

Saeed, John. (1993) *Somali Reference Grammar*. Kensington, MD: Dunwoody Press.

Saeed, John. (1999) *Somali*. Amsterdam: Benjamins.

Saffran, Jenny R., Richard N. Aslin and Elissa L. Newport. (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Sagart, Laurent, Pierre Hallé, Bénédicte de Boysson-Bardies, and Catherine Arabia-Guidet. (1986) Tone production in Modern Standard Chinese: An electromyographic investigation. *Cahiers de linguistique Asie Orientale* 15, 205–221.

Sagey, Elizabeth. (1986) *The Representation of Features and Relations in Nonlinear Phonology*. MIT dissertation. Published by Garland Publishing, New York, 1991.

Samek-Lodovici, Vieri. (2005) Prosody syntax interaction in the expression of focus. *Natural Language and Linguistic Theory* 23, 687–755.

Sample, Ward. (1976) *The Application of Rules in the Phonology of Olukisa*. Indiana University dissertation.

Sampson, Geoffrey. (1975) One fact needs one explanation. *Lingua* 36, 231–239.

Samuels, B. (2009) *The Structure of Phonological Theory*. Harvard University dissertation.

Sanders, Gerald. (1972a) *Equational Grammar*. The Hague: Mouton.

Sanders, Gerald. (1972b) *The Simplex Feature Hypothesis*. Bloomington, IN: Indiana University Linguistics Club.

Sanders, Gerald. (1974) Precedence relations in language. *Foundations of Language* 11, 361–400.

Sanders, Nathan. (1999) Same-edge alignment with opposite-edge effects, in *Proceedings of the 18th Annual Meeting of the West Coast Conference on Formal Linguistics*.

Sanders, Nathan. (2000) Intra-representational correspondence and the realization of empty morphemes. Master's thesis, University of Southern California.

Sanders, Nathan. (2003) *Opacity and Sound Change in the Polish Lexicon*. University of California, Santa Cruz dissertation.

Sandler, Wendy. (1989) *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht: Foris.

Sandler, Wendy. (1993) A sonority cycle in American Sign Language. *Phonology* 10, 2, 243–279.

Sandler, Wendy. (1999) Cliticization and prosodic words in a sign language, in *Studies on the Phonological Word*, ed. by T. Alan Hall and U. Kleinhenz, 223–255, Amsterdam: Benjamins.

Sandler, Wendy and Diane Lillo-Martin. (2006) *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sandler, Wendy, Irit Meir, Carol Padden, and Mark Aronoff. (2005) The emergence of grammar in a new sign language, in *Proceedings of the National Academy of Sciences*, 102(7), 2661–2665.

Sankoff, David, Sali Tagliamonte, and Eric Smith. (2005) Goldvarb X: a variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto. Available at http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm.

Santa Ana, Otto. (1991) *Phonetic Simplification Processes in the English of the Barrio: A Cross-Generational Sociolinguistic Study of the Chicagos of Los Angeles*. University of Pennsylvania dissertation.

Santa Ana, Otto. (1992) Chicago English evidence for the exponential hypothesis: A variable rule pervades lexical phonology. *Language Change and Variation* 4, 275–288.

Sapir, David. (1965) *A Grammar of Diola-Fogny.* Cambridge: Cambridge University Press.

Sapir, Edward. (1915) Abnormal types of speech in Nootka. Technical report, Canadian Department of Mines Geological Survey Memoir 62, Anthropological Series 5.

Sapir, Edward. (1933) La réalité psychologique des phonèmes. *Journal de Psychologie Normale et Pathologique* 30, 247–265.

Sapir, Edward and Morris Swadesh. (1960) *Yana Dictionary*. Berkeley, CA: University of California Press.

Sapir, Edward and Morris Swadesh. (1978) *Nootka Texts: Tales and Ethnological Narratives, with Grammatical Notes and Lexical Material*. New York: AMS Press.

Saporta, Sol. (1955) Frequency of consonant clusters. *Language* 31, 25–30.

Saraçlar, Murat and Sanjeer Khudanpur. (2004) Pronunciation change in conversational speech and its implications for automatic speech recognition. *Computer Speech and Language* 18, 375–395.

Sasa,Tomomasa. (2004) Directionality in harmony: Evidence from Pulaar. Master's thesis, University of Southern California.

Sasa, Tomomasa. (2006) Directionality in harmony: Evidence from Pulaar. Presented at the 14th Manchester Phonology Meeting, Manchester UK.

Saussure, Ferdinand de. (1916) *Cours de linguistique générate.* Paris: Payot.

Saussure, Ferdinand de. (1916/1959) *Course in General Linguistics.* New York: McGraw-Hill.

Saussure, Ferdinand de. (1995) *Phonétique: Il manoscritto di Harvard Houghton Library bMS Fr*, 266(8). Florence: Unipress.

Savà, Graziano. (2005) *A Grammar of Ts'amakko.* Cologne: Rüdiger Köppe Verlag.

Scarborough, Rebecca. (2004) *Coarticulation and the Structure of the Lexicon*. UCLA dissertation.

Schachter, Paul and Victoria Fromkin. (1968) *A Phonology of Akan: Akuapem, Asante, Fante*, volume 9 of *Working Papers in Phonetics*. Los Angeles: University of California Press.

Schachter, Paul and Fe T. Otanes. (1972) *Tagalog Reference Grammar*. Berkeley, CA: University of California Press.

Schacter, Daniel L. and Barbara A. Church. (1992) Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 915–930.

Schadeberg, Thilo C. (1973) Kinga: A restricted tone language. *Studies in African Linguistics* 4, 23–48.

Schane, Sanford A. (1965) *Phonological and Morphological Structure of French*. MIT dissertation.

Schane, Sanford A. (1971) The phoneme revisited. *Language* 47, 503–521.
Schane, Sanford A. (1973) *Generative Phonology*. Englewood Cliffs, NJ: Prentice-Hall.
Schane, Sanford A. (1984) Fundamentals of particle phonology. *Phonology Yearbook* 1, 129–155.
Schane, Sanford A. (1990) Lowered height, laxness, and Retracted Tongue Root: Different manifestations of phonological aperture. *Word* 41, 1–16.
Schane, Sanford A. (1995) Diphthongisation in particle phonology, in *Handbook of Phonological Theory*, ed. by John A. Goldsmith, 586–605, Oxford: Blackwell.
Scheer, Tobias. (2004) *A Lateral Theory of Phonology. What is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.
Scheer, Tobias. (2007) On the status of word-initial clusters in Slavic (and elsewhere), in *Annual Workshop on Formal Approaches to Slavic Linguistics. The Toronto Meeting 2006*, ed. by Richard Compton, Magdalena Goledzinowska, and Ulyana Savchenko, 346–364, Ann Arbor, MI: Michigan Slavic Publications.
Scheibman, Joanna and Joan Bybee. (1999) The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37, 575–596.
Schein, Barry and Donca Steriade. (1986) On geminates. *Linguistic Inquiry* 17, 691–744.
Schick, Brenda. (1987) *The Acquisition of Classifier Predicates in American Sign Language*. Purdue University dissertation.
Schmidt, Deborah. (1992) Compensatory lengthening in a segmental moraic theory of representation. *Linguistics* 30, 513–534.
Scholes, Robert J. (1966) *Phonotactic Grammaticality*. The Hague: Mouton.
Scholz, Barbara C. and Geoffrey K. Pullum. (2007) Tracking the origins of transformational generative grammar. *Journal of Linguistics* 43, 701–723.
Schourup, Lawrence. (1972) Characteristics of vowel nasalization, in *Papers in Linguistics*, volume 5, 530–548.
Schouten, M. E. H. (1982) T-deletie in de stad Utrecht: schoolkinderen en grootouders. *Forum der Lettern* 23, 282–291.
Schouten, M. E. H. (1984) T-deletie in het zuiden van die provincie Utrecht. *Taal en Tongval* 36, 162–173.
Schuppler, Barbara, Mirjam Ernestus, Odette Scharenborg, and Lou Boves. (2011) Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics* 39, 96–109.
Schütz, Albert. (1978) English loanwords in Fijian, in *Fijian Language Studies: Borrowing and Pidginization*, ed. by Albert Schütz, volume 4, 2–50, Fiji: Fiji Museum.
Schütz, Albert. (1985) *The Fijian Language*. Honolulu, HI: University of Hawaii Press.
Schütze, Carson. (1996) *The Empirical Base of Linguistics*. Chicago: University of Chicago Press.
Schütze, Carson. (2005) Thinking about what we are asking speakers to do, in *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, ed. by S. Kepser and M. Reis, 457–484, The Hague: Mouton de Gruyter.
Schweickert, Richard. (1993) A multinomial processing tree model of degradation and redintegration in immediate recall. *Memory and Cognition* 21, 168–175.
Scobbie, James M. (1997) *Autosegmental Representation in a Declarative Constraint-Based Framework*. New York: Garland Publishing.
Scobbie, James M., John S. Coleman, and Steven Bird. (1996) Key aspects of declarative phonology, in *Current Trends in Phonology: Models and Methods*, ed.

by Jacques Durand and Bernard Laks, volume 2, 685–710, European Studies Research Institute (ESRI) and Manchester: University of Salford.

Scobbie, James M., Fiona Gibbon, William J. Hardcastle, and Paul Fletcher. (2000) Covert contrast as a stage in the acquisition of phonetics and phonology, in *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*, ed. by M. Broe and J. Pierrehumbert, 194–207, Cambridge: Cambridge University Press.

Segui, Juan, Ulricht Frauenfelder, and Pierre Hallé. (2001) Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing, in *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler,* ed. by Emmanuel Dupoux, 195–208, Cambridge, MA: MIT Press.

Seidl, Amanda. (2001) *Minimal Indirect Reference: A Theory of the Syntax-Phonology Interface*. London, New York: Routledge.

Seidl, Amanda and Eugene Buckley. (2005) On the learning of arbitrary phonological rules. *Language Learning and Development* 1, 289–316.

Seidl, Amanda, Elizabeth Johnson, A. Redman, and Diane Brentari. (2004) Segmentation of clauses in English, Dutch, and American Sign Language. Presented at the Boston Child Language Conference.

Seifart, Frank. (2005) *The Structure and Use of Shape-Based Noun Classes in Miraña (North West Amazon)*. Radbound Universiteit Nijmegen dissertation.

Sejnowski, Terrence J. and Charles R. Rosenberg. (1987) NETtalk: A parallel network that learns to pronounce English text. *Complex Systems* 1, 145–168.

Selkirk, Elisabeth O. (1972) *The Phrase Phonology of English and French.* MIT dissertation. Published by Garland Publishing, New York, 1980.

Selkirk, Elisabeth O. (1974) French liaison and the X-bar convention. *Linguistic Inquiry* 5, 573–590.

Selkirk, Elisabeth O. (1978) On prosodic structure and its relation to syntactic structure, Indiana University Linguistics Club (IULC).

Selkirk, Elisabeth O. (1980a) Prosodic domains in phonology: Sanskrit revisited, in *Juncture*, ed. by Mark Aronoff and Mary-Louise Kean, Saratoga: Anma Libri.

Selkirk, Elisabeth O. (1980b) The role of prosodic categories in English word stress. *Linguistic Inquiry* 11, 563–605.

Selkirk, Elisabeth O. (1981a) The nature of phonological representation, in *The Cognitive Representation of Speech*, ed. by John M. Anderson, John Laver, and Terry Myers, 379–388, Dordrecht: North Holland Publishing Company.

Selkirk, Elisabeth O. (1981b) On prosodic structure and its relation to syntactic structure, in *Nordic Prosody II: Papers from a Symposium on the Proceedings of the Second Symposium on the Prosody of the Nordic Languages, Trondheim, 1980*, ed. by Thorstein Fretheim, 111–140, Trondheim. Distributed by Indiana University Linguistics Club.

Selkirk, Elisabeth O. (1982) The syllable, in *The Structure of Phonological Representations*, ed. by Harry van der Hulst and Norval Smith, volume 2, 337–383, Dordrecht: Foris.

Selkirk, Elisabeth O. (1984a) On the major class features and syllable theory, in *Language Sound Structure: Studies in Phonology Dedicated to Morris Halle by his Teacher and Students*, ed. by Mark Aronoff and Richard T. Oehrle, 107–113, Cambridge, MA: MIT Press.

Selkirk, Elisabeth O. (1984b) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth O. (1986) On derived domains in sentence phonology. *Phonology Yearbook* 3, 371–405.

Selkirk, Elisabeth O. (1990) The two-root node theory of length, in *University of Massachusetts Occasional Papers 14*, 123–171.

Selkirk, Elisabeth O. (1995) Sentence prosody: Intonation, stress and phrasing, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 550–569, Oxford: Blackwell.

Selkirk, Elisabeth O. (1996) The prosodic structure of function words, in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, ed. by James L. Morgan and Katherine Demuth, 187–213, Mahwah, NJ: Lawrence Erlbaum.

Selkirk, Elisabeth O. (2000) The interaction of constraints on prosodic phrasing, in *Prosody: Theory and Experiments*, ed. by Merle Home, 231–262, Dordrecht: Kluwer.

Selkirk, Elisabeth O. (2002) Contrastive focus vs. presentational focus: Prosodic evidence from Right Node Raising in English, in *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody*, ed. by Isabella Marlien and Bernard Bel, 643–646, Aix-en-Provence, France.

Selkirk, Elisabeth O. (2005) Comments on Intonational Phrasing, in *Prosodies*, ed. by Sónia Frota, Marina Vigário, and Maria João Freitas, 11–58, Berlin: Mouton de Gruyter.

Selkirk, Elisabeth O. (2006) Strong minimalist spell-out of prosodic phrases. Talk presented at GLOW Workshop on Phonological Phrasing, Barcelona, April 2006.

Selkirk, Elisabeth O. (2007) Focus prominence drives focus phrasing: Bengali intonation revisited, in *Topic and Focus: Cross-linguistic Perspectives on Intonation and Meaning*, ed. by Chung-min Lee, Matthew Gordon, and Daniel Buring, 217–246, Dordrecht: Springer.

Selkirk, Elisabeth O. (2008) Contrastive focus, givenness and the unmarked status "discourse-new." *Acta Linguistia Hungarica* 55, 331–346.

Selkirk, Elisabeth O. (2009) On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136, 35–74.

Selkirk, Elisabeth O. and Tong Shen. (1990) Prosodic domains in Shanghai Chinese, in *The Phonology-Syntax Connection*, ed. by Sharon Inkelas and Draga Zec, 313–337, Chicago: University of Chicago Press.

Selkirk, Elisabeth O., Takahito Shinya, and Shigeto Kawahara. (2004) Phonology and phonetic effects of phrase length on F0 in Japanese, in *Proceedings of the Second International Conference on Speech Prosody*, ed. by Bernard Bel and Isabelle Marlien, 183–186, Japan: Nara.

Selkirk, Elisabeth O., Takahito Shinya, and Mariko Sugahara. (2003) Degree of lowering in Japanese as a reflex of prosodic structure organization, in *Proceedings of the 15th International Congress of Phonetic Sciences*, 491–494, Barcelona.

Selkirk, Elisabeth O. and Koichi Tateishi. (1988) Constraints on minor phrase formation in Japanese, in *Proceedings of the 24th Annual Meeting of the Chicago Linguistics Society*, ed. by Gary Larson and Diane Brentari, 316–336, Chicago, IL: Chicago Linguistics Society.

Selkirk, Elisabeth O. and Koichi Tateishi. (1991) Syntax and downstep in Japanese, in *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*, ed. by Carol Georgopoulos and Roberta Ishihara, 519–543, Dordrecht: Kluwer.

Semon, Richard W. (1923) *Mnemische Empfindungen.* London: Allen and Unwin.

Senghas, Ann. (1995) *Children's Contribution to the Birth of Nicaraguan Sign Language*. MIT dissertation.

Senghas, Ann and Marie Coppola. (2001) Children creating language: How

Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science* 12, 323–328.

Senn, Alfred. (1966) *Handbuch der Litauischen Sprache*. Heidelberg: Carl Winter.

Sereno, Joan and Allard Jongman. (1997) Processing of English inflectional morphology. *Memory and Cognition* 25, 425–437.

Sevald, Christine A., Gary S. Dell, and Jennifer S. Cole. (1995) Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language* 34, 807–820.

Sezer, Engin. (1981) On non-final stress in Turkish. *Journal of Turkish Studies* 5, 61–69.

Sezer, Engin. (1986) An autosegmental analysis of compensatory lengthening in Turkish, in *Studies in Compensatory Lengthening*, ed. by Leo Wetzels and Engin Sezer, 221–251, Dordrecht: Foris.

Shademan, Shabnam. (2006) Is phonotactic knowledge grammatical knowledge?, in *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, and Michael Scanlon, 371–379, Somerville, MA: Cascadilla Press.

Shademan, Shabnam. (2007) *Grammar and Analogy in Phonotactic Well-Formedness*. UCLA dissertation.

Shahin, Kimary. (2002) *Postvelar Harmony*. Amsterdam: Benjamins.

Shannon, Claude E. (1948) A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423.

Shattuck-Hufnagel, Stephanie, Katherine Demuth, Helen M. Hanson, and Kenneth N. Stevens. (in press) Acoustic cues to stop-coda voicing contrasts in the speech of American English 2–3-year-olds, in *Where do Feature Come from? The Nature and Sources of Phonological Primitives*, ed. by G. N. Clements and R. Ridouane, Amsterdam: North-Holland Linguistic Series, Elsevier.

Shattuck-Hufnagel, Stephanie and Dennis H. Klatt. (1979) The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18, 41–55.

Shattuck-Hufnagel, Stephanie and Alice Turk. (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25, 193–247.

Shatzman, Keren and René Kager. (2007) A role for phonotactic constraints in speech perception, in *Proceedings of the 16th International Congress of Phonetic Sciences*, ed. by J. Trouvain and W. J. Barry, 1409–1412, Dudweiler, Germany: Pirrot.

Shaw, Patricia A. (1985) Modularisation and substantive constraints in Dakota lexical phonology. *Phonology* 2, 173–202.

Shaw, Patricia A. (1991) Consonant harmony systems: The special status of coronal harmony, in *The Special Status of Coronals: Internal and External Evidence*, ed. by Carole Paradis and Jean-François Prunet, 125–157, San Diego, CA: Academic Press.

Shaw, Patricia A. (1993) The prosodic constituency in minor syllables, in *The Proceedings of the 12th West Coast Conference on Formal Linguistics*, ed. by E. Duncan, D. Farkas, and P. Spaelti, 117–132.

Shaw, Patricia A. (2009) Inside access, in *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, Cambridge, MA: MIT Press.

Shepherd, Michael A. Constraint. (2003) interactions in Spanish phonotactics: An Optimality Theory analysis of syllable-level phenomena in the Spanish language. Master's thesis, California State University, Northridge. ROA-639.

Sherer, Timothy. (1994) *Prosodic Phonotactics*. University of Massachusetts dissertation.

Sherzer, Joel. (1970) Talking backwards in Cuna: The sociological reality of phonological descriptions. *Southwestern Journal of Anthropology* 26, 343–353.

Shibatani, Masayoshi. (1973) The role of surface phonetic constraints in generative phonology. *Language* 49, 87–106.

Shih, Chilin and Greg Kochanski. (2000) Chinese tone modeling with Stem-ML, in *Proceedings of the Sixth International Conference on Spoken Language Processing*, volume 2, 67–70.

Shinohara, Shigeko. (1997) *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Paris: Université de la Sorbonne Nouvelle, Paris III, dissertation.

Shintel, Hadas and Howard Nussbaum. (2007) The sound of motion in spoken language: Visual information conveyed in acoustic properties of speech. *Cognition* 105, 681–690.

Shintel, Hadas, Howard Nussbaum, and Arika Okrent. (2006) Analog acoustic expression in speech communication. *Journal of Memory and Language* 55, 167–177.

Shinya, Takahito, Elisabeth O. Selkirk, and Shigeto Kawahara. (2004) Rhythmic boost and recursive minor phrase in Japanese, in *Proceedings of the 2nd International Conference on Speech Prosody*, 183–186, Nara, Japan.

Sibanda, Galen. (2004) *Verbal Phonology and Morphology of Ndebele*. University of California, Berkeley dissertation.

Sievers, Eduard. (1901) *Grundzüge der Phonetik*. Leipzig: Breitkopf und Hartel, 5th edition.

Silverman, Daniel. (1992) Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9, 289–328.

Silverman, Daniel. (1996) Phonology at the interface of phonetics and morphology: Root-final laryngeals in Chong, Korean, and Sanskrit. *Journal of East Asian Linguistics* 5, 301–322.

Silverman, Daniel. (2004) On the phonetic and cognitive nature of alveolar stop allophony in American English. *Cognitive Linguistics* 15, 69–93.

Silverman, Daniel. (2006a) *A Critical Introduction to Phonology: of Sound, Mind, and Body*. London: Continuum.

Silverman, Daniel. (2006b) The diachrony of labiality in Trique, and the functional relevance of gradience and variation, in *Laboratory Phonology VIII: Varieties of Phonological Competence*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 135–154, Berlin: Mouton de Gruyter.

Silverman, Daniel. (2010) Neutralization and anti-homophony in Korean. *Journal of Linguistics* 46, 453–482.

Silverman, Kim. (1986) F0 cues depend on intonation: The case of the rise after voiced stops. *Phonetica* 43, 76–92.

Simonet, Miquel, Marcos Rohena-Madrazo, and Mercedes Paz. (2008) Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish, in *Selected Proceedings of the 3rd, Conference on Laboratory Approaches to Spanish Phonology*, ed. by Laura Colantoni and Jeffrey Steele, 72–86, Somerville, MA: Cascadilla Proceedings Project.

Singh, Rajendra. (1987) Well-formedness conditions and phonological theory, in *Phonologica, 1984,* ed. by Wolfgang Dressier, H. C. Luschützky, O. E. Pfeiffer, and J. R. Rennison, 273–285, Cambridge: Cambridge University Press.

Singh, Rajendra. (1996) Natural phono(morpho)logy: A view from the outside, in *Natural Phonology: The State of the Art*, ed. by Bernhard Hurch and Richard Rhodes, 1–38, Berlin: Mouton.

Singler, John. (1980) The status of lexical associations and the Obligatory Contour Principle in the analysis of tone languages, in *Proceedings of the 6th Annual Meeting of the Berkeley Linguistics Society*, 442–456.

Siraj, Pasha. (2008) Intonation in Singaporean English is stress-dependent. Poster at Tone and Intonation in Europe 3 (Lisbon).

Sjölander, Kare and Jonas Beskow. (2000) WaveSurfer – an open-source speech tool, in *Proceedings of the 6th International Conference on Spoken Language Processing*, volume 4, 464–467.

Skinner, B. F. (1957) *Verb Behavior*. New York: Appleton-Century-Crofts.

Skoruppa, Katrin and Sharon Peperkamp. (2008) Decomposing naturalness in phonological rule learning: The role of phonetic distance, in *Proceedings of the 39th Annual Meeting of the North East Linguistic Society*, volume 39.

Skousen, Royal. (1989) *Analogical Modeling of Language*. Dordrecht: Kluwer.

Skousen, Royal. (2002) Analogical modeling and quantum computing, in *Analogical Modeling: An Exemplar-based Approach to Language*, ed. by R. Skousen, D. Lonsdale, and D. B. Parkinson, 319–346, Amsterdam: Benjamins.

Sloat, Clarence. (1966) *Phonological Redundancy Rules in Coeur d'Alene*. University of Washington dissertation.

Smiljanić, Rajka. (2006) Early vs. late focus: pitch-peak alignment in two dialects of Serbian and Croatian, in *Laboratory Phonology VIII*, ed. by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 394–518, The Hague: Mouton de Gruyter.

Smith, Caroline. (1992) *The Temporal Organization of Vowels and Consonants*. Yale University dissertation.

Smith, Jennifer L. (1999) Noun faithfulness and accent in Fukuoka Japanese, in *Proceedings of the 18th Annual Meeting of the West Coast Conference on Formal Linguistics*, ed. by Sonya Bird, Andrew Carnie, Jason Haugen, and Peter Norquest, 519–531, Somerville, MA: Cascadilla Press.

Smith, Jennifer L. (2000a) Positional faithfulness and learnability in Optimality Theory, in *Proceedings of ESCOL '99*, ed. by Rebecca Daly and Anastasia Riehl, 203–214, Ithaca, NY: CLC Publications.

Smith, Jennifer L. (2002) *Phonological Augmentation in Prominent Positions*. University of Massachusetts, Amherst dissertation.

Smith, Jennifer L. (2003) Onset sonority constraints and subsyllabic structure. University of North Carolina, manuscript.

Smith, Jennifer L. (2004) Making constraints positional: Toward a compositional model of CON. *Lingua* 114, 1433–1464.

Smith, Jennifer L. (2005) *Phonological Augmentation in Prominent Positions*. London: Routledge.

Smith, N. V. (1969) The Nupe verb. *African Language Studies* 10, 90–160.

Smith, N. V. (1973) *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.

Smith, Norval. (2000) Dependency phonology meets OT: A proposal for a new approach to segmental structure, in *Optimality Theory: Phonology, Syntax, and Acquisition*, ed. by Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 234–278, Oxford: Oxford University Press.

Smolensky, Paul. (1986) Information processing in dynamical systems: Foundations of harmony theory, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, ed. by David Rumelhart, James McClelland, and the PDP Research Group, 194–281, Cambridge, MA: MIT Press/Bradford Books.

Smolensky, Paul. (1993) Harmony, markedness and phonological activity. Presented at the Rutgers Optimality Workshop 1.

Smolensky, Paul. (1996a) The initial state and "richness of the base" in Optimality Theory. Technical report, Johns Hopkins University.

Smolensky, Paul. (1996b) On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27, 720–731.

Smolensky, Paul. (1997) Constraint interaction in generative grammar II: Local conjunction. Presented at the Hopkins Optimality Theory Workshop/ Maryland Mayfest 1997.

Smolensky, Paul. (2006) Computational levels and integrated connectionist/ symbolic explanation, in *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar (Vol. 2, Linguistic and Philosophical Implications)*, ed. by Paul Smolensky and Géraldine Legendre, 503–592, Cambridge, MA: MIT Press.

Smolensky, Paul and Géraldine Legendre. (2006) *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press/ Bradford Books.

Smolensky, Paul, Géraldine Legendre, and Bruce Tesar. (2006) Optimality Theory: The structure, use, and acquisition of grammatical knowledge, in *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar (Vol. 1: Cognitive Architecture)*, ed. by Paul Smolensky and Géraldine Legendre, 453–544, Cambridge MA: MIT Press.

Snider, Keith. (1999) *The Geometry and Features of Tone*. Dallas: University of Texas at Arlington.

So, Lydia K. H. (1996) Tonal changes in Hong Kong Cantonese. *Current Issues in Language and Society* 3, 186–189.

Soanes, Catherine and Angus Stevenson (eds.) (2004) *Concise Oxford English Dictionary*. Oxford University Press, 11th edition.

Söderpalm, Ewa. (1979) *Speech Errors in Normal and Pathological Speech (Travaux de l'Institute de Linguistique de Lund, 14)*. Lund: CWK Gleerup.

Sohn, H. S. (1987) *Underspecification in Korean Phonology*. University of Illinois dissertation.

Solé, Maria-Josep. (2007) Controlled and mechanical properties in speech: A review of the literature, in *Experimental Approaches to Phonology*, ed. by Maria-Josep Solé, Patrice S. Beddor, and Manjari Ohala, 302–321, Oxford: Oxford University Press.

Sommerstein, Alan H. (1974) On phonotactically motivated rules. *Journal of Linguistics* 19, 71–94.

Sommerstein, Alan H. (1977) *Modern Phonology*. London: Edward Arnold.

Song, Jae Yung and Katherine Demuth. (2008) Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words. *Language and Speech* 51, 385–402.

Song, Jae Yung, Megha Sundara, and Katherine Demuth. (2009) Phonological constraints on children's production of English third person singular *-s*. *Journal of Speech, Language, and Hearing Research* 52, 623–642.

Spa, Jaap. (1973) *Traits et tons en enya: phonologie générative d'une langue bantoue*. Number 20 in Archives d'Anthropologic, Tervueren, Musée Royal de l'Afrique Centrale.

Spencer, Andrew. (1986) Towards a theory of phonological development. *Lingua*, 68, 3–38.

Spencer, Andrew. (1998) Morphophonological operations, in *Handbook of Morphology*, ed. by Andrew Spencer and Arnold Zwicky, 123–143, Oxford: Blackwell.

Sproat, Richard. (1986) Malayalam compounding: A non-stratum ordered account, in *Proceedings of the 5th West Coast Conference on Formal Linguistics*, ed. by Mary Dalrymple *et al.*, 268–288, Stanford, CA: Stanford Linguistics Association.

Stallcup, Kenneth. (1980) Noun classes in Esimbi, in *Noun Classes in the Grassfields Bantu Borderland*, ed. by Larry M. Hyman, 139–153, Los Angeles: University of Southern California.

Stampe, David. (1969) The acquisition of phonetic representation. Presented at the 5th Regional Meeting of the Chicago Linguistic Society.

Stampe, David. (1972) *A Dissertation on Natural Phonology*. New York: University of Chicago dissertation. Published by Garland Publishing, New York, 1979.

Stanley, Richard. (1967) Redundancy rules in phonology. *Language* 43, 393–436.

Staubs, Robert, Michael Becker, Christopher Potts, Patrick Pratt, John J. McCarthy, and Joe Pater. (2010) *OT-Help 2.0*. Software package. Amherst, MA: University of Massachusetts, Amherst.

Stell, Nélida Noemi. (1972) *Fonología de la lengua Alulaj*. Buenos Aires: Universidad de Buenos Aires.

Stemberger, Joseph P. (2004) Neighbourhood effects on error rates in speech production. *Brain and Language* 90, 413–422.

Stemberger, Joseph P. and Barbara H. Bernhardt. (1997) Phonological constraints and morphological development, in *Proceedings of the 21st Annual Boston University Conference on Language Development*, ed. by E. Hughes, M. Hughes, and A. Greenhill, 602–614, Somerville, MA: Cascadilla Press.

Stemberger, Joseph P. and Brian MacWhinney. (1988) Are lexical forms stored in the lexicon?, in *Theoretical Morphology: Approaches in Modern Linguistics*, ed. by M. Hammond and M. Noonan, 101–116, London: Academic Press.

Stemberger, Joseph P. and Carol Stoel-Gammon. (1991) The underspecification of coronals: Evidence from language acquisition and performance errors, in *The Special Status of Coronals: Internal and External Evidence*, ed. by Carole Paradis and Jean-François Prunet, 181–200, San Diego, CA: Academic Press.

Stemberger, Joseph P. and Rebecca Treiman. (1986) The internal structure of word-initial consonant clusters. *Journal of Memory and Language* 25, 163–180.

Stenneken, Prisca, Roelien Bastiaanse, Walter Huber and Arthur M. Jacobs. (2005) Syllable structure and sonority in language inventory and aphasic neologisms. *Brain and Language* 95 (2), 280–292.

Steriade, Donca. (1982) *Greek Prosodies and the Nature of Syllabification*. MIT dissertation.

Steriade, Donca. (1986) A Note on Coronal. Manuscript, MIT.

Steriade, Donca. (1987a) Locality conditions and feature geometry, in *Proceedings of the 17th Annual Meeting of the North East Linguistic Society*, ed. by J. McDonough and B. Plunkett, 595–617.

Steriade, Donca. (1987b) Redundant values, in *Papers from the Chicago Linguistic Society Parasession on Autosegmental and Metrical Phonology*, ed. by Anna Bosch, Barbara Need, and Eric Schiller, 339–362.

Steriade, Donca. (1988) Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology* 5, 73–155.

Steriade, Donca. (1991) Moras and other slots. *Formal Linguistic Society of Mid-America* 1, 254–280.

Steriade, Donca. (1994) Positional neutralization and the expression of contrast. Unpublished manuscript, UCLA.

Steriade, Donca. (1995) Underspecification and Markedness, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 114–174, Oxford: Blackwell.

Steriade, Donca. (1999a) Alternatives to syllable-based accounts of consonantal phonotactics, in *Proceedings of the 2nd Linguistics and Phonetics Conference*, ed. by Osamu Fujimura and B. Palek, 205–245, Prague: Charles University Press.

Steriade, Donca. (1999b) Phonetics in phonology: The case of laryngeal neutralization, in *Papers in phonology*, ed. by Matthew Gordon, volume 3, 25–146, Los Angeles: Department of Linguistics, UCLA.

Steriade, Donca. (2000) Paradigm uniformity and the phonetics/phonology boundary, in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, ed. by Michael B. Broe and Janet B. Pierrehumbert, Cambridge: Cambridge University Press.

Steriade, Donca. (2001) Directional asymmetries in place assimilation, in *The Role of Speech Perception in Phonology*, ed. by Elizabeth V. Hume and Keith Johnson, 219–250, San Diego, CA: Academic Press.

Steriade, Donca. (2009) The phonology of perceptibility effects: The P-map and its consequences for constraint organization, in *The Nature of the Word: Studies in Honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 151–179, Cambridge, MA: MIT Press.

Stevens, Kenneth N. (1989) On the quantal nature of speech. *Journal of Phonetics* 17, 3–45.

Stevens, Kenneth N. and Shiela Blumstein. (1981) The search for invariant acoustic correlates of phonetic feature, in *Perspectives on the Study of Speech*, ed. by Peter D. Eimas and Joanne L. Miller, 1–38, Hillsdale, NJ: Lawrence Erlbaum.

Stevens, Kenneth N. and Samuel Jay Keyser. (1989) Primary features and their enhancement in consonants. *Language* 65, 81–106.

Stevens, Kenneth N. and Samuel Jay Keyser. (2010) Quantal theory, enhancement, and overlap. *Journal of Phonetics* 38, 10–19.

Stevens, Kenneth N., Samuel Jay Keyser, and H. Kawasaki. (1986) Toward a phonetic and phonological theory of redundant features, in *Invariance and Variability in Speech Processes*, ed. by J. Perkell and D. Klatt, 426–449, Hillsdale, NJ: Lawrence Erlbaum.

Stevick, E. W. (1969) Pitch and duration in Ganda. *Journal of African Languages* 8, 1–28.

Stewart, J. M. (1967) Tongue root position in Akan vowel harmony. *Phonetica* 16.

Stites, Jessica, Katherine Demuth and Cecilia Kirk. (2004) Markedness vs. frequency effects in coda acquisition, in *Proceedings of the Boston University Conference on Language Development*, volume 28, 565–576, Somerville, MA: Cascadilla Press.

Stockall, L., A. Stringfellow, and Alec Marantz. (2004) The precise timecourse of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. *Brain and Language* 90, 88–94.

Stoel-Gammon, Carol and Eugene H. Buder. (1999) Vowel length, postvocalic voicing and VOT in the speech of two-year olds, in *Proceedings of the 13th International Conference of Phonetic Sciences*, volume 3, 2485–2488.

Stoel-Gammon, Carol and Eugene H. Buder. (2002) American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *The Journal of the Acoustical Society of America* 111, 1854–1864.

Stoel-Gammon, Carol and Joseph P. Stemberger. (1994) Consonant harmony and phonological underspecification in child speech, in *First and Second Language Phonology,* ed. by M. S. Yavas, 63–80, San Diego, CA: Singular Publishing Group.

Stokoe, William. (1960) Sign language structure: An outline of the visual communication systems of the American Deaf, in *Studies in Linguistics*, *Occasional Papers*, Silver Spring, MD: Linstok Press.

Stokoe, William, Dorothy Casterline, and Carl Croneberg. (1965) *A Dictionary of American Sign Language on Linguistic Principles*. Silver Spring, MD: Linstok Press.

Stoll, R. P. (1955) *La tonétique des langues Bantu et semi-Bantu du Cameroun*. Yaounde: IFAN.

Stone, Maureen and S. Hamlet. (1982) Variation in jaw and tongue gestures observed during the production of

unstressed /d/s and flaps. *Journal of Phonetics* 10, 401–415.
Storkel, H. L. (2004) The emerging lexicon of children with phonological delays: Phonotactic constraints and probability in acquisition. *Journal of Speech, Language, and Hearing Research* 47, 1194–1212.
Straus, Oliver H. (1950) The relation of phonetics and linguistics to communication theory. *Journal of the Acoustical Society of America* 22, 709–711.
Struijke, Caro. (2000) *Existential Faithfulness: A Study of Reduplicative TETU, Feature Movement, and Dissimilation*. University of Maryland dissertation.
Stump, Gregory T. (1991) On the theoretical status of position class restrictions, in *Yearbook of Morphology 1991*, ed. by Geert Booij and Jaap van Marle, 211–241, Dordrecht: Kluwer.
Stump, Gregory T. (2001) *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
Sugito, Miyoko. (2003) Timing relationships between prosodic and segmental control in Osaka Japanese word accent. *Phonetica* 60, 1–16.
Sun, Jiping and Li Deng. (2002) An overlapping-feature-based phonological model incorporating linguistic constraints: Applications to speech recognition. *Journal of the Acoustical Society of America*, 111, 1086–1101.
Sundara, Megha, Katherine Demuth, and Patricia Kuhl. (2011) Sentence-position effects on children's comprehension and production of English third person singular *-s*. *Journal of Speech, Language, and Hearing Research* 54, 55–71.
Suomi, Kari. (1983) Palatal harmony: A perceptually motivated phenomenon? *Nordic Journal of Linguistics* 6, 1–35.
Suomi, Kari, James McQueen, and Anne Cutler. (1997) Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language* 36, 422–444.
Sussman, Harvey M., Helen A. McCaffrey, and Sandra A. Matthews. (1991) An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90, 1309–1325.
Svantesson, Jan-Olof. (1985) Vowel harmony shift in Mongolian. *Lingua*, 67, 283–327.
Svantesson, Jan-Olof, Anna Tsendina, Anastasia Karlsson, and Vivan Franzén. (2005) *The Phonology of Mongolian*. Oxford: Oxford University Press.
Swadesh, Morris. (1946) Chitimacha, in *Linguistic Structures of Native America*, ed. by Cornelius Osgood, 312–336, New York: Viking Fund Publications in Anthropology.
Swick, Joyce. (1966) Chuave phonological hierarchy. *Pacific Linguistics Series A* 7, 33–148.
Swingley, Dan. (2007) Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology* 43, 454–464.
Sybesma, Rint and Boya Li. (2007) The dissection and structural mapping of Cantonese sentence final particles. *Lingua* 117, 1739–1783.
Szendroi, Kriszta. (2001) *Focus and the Syntax-Phonology Interface*. University College London dissertation.
Szigetvári, Péter. (1999) *VC Phonology: A Theory of Consonant Lenition and Phonotactics*. Eötvös Loránd University dissertation.
Szigetvári, Péter. (2000) Why CVCV. *The Even Yearbook* 4, 117–152. ELTE SEAS Working Papers in Linguistics, Budapest.
Szigetvári, Péter and Tobias Scheer. (2005) Unified representations for stress and the syllable. *Phonology* 22, 37–75.
Tabain, Marija and Kristine Rickard. (2007) A preliminary EPG study of stop consonants in Arrernte, in *Proceedings of International Congress of Phonetic Sciences.*
Tadadjeu, Maurice. (1974) Floating tones, shifting rules, and downstep in

Dschang-Bamileke. *Studies in African Linguistics, Supplement 5*, 283–290.
Takahashi, Toyomi. (1993) A farewell to constituency. *UCL Working Papers in Linguistics* 5, 375–410.
Takahashi, Toyomi. (2004) *Syllable Theory without Syllables*. SOAS London dissertation.
Tang, Gladys, Diane Brentari, Felix Sze, and Carolina González. (2010) Crosslinguistic patterns of prosodic cues in sign languages: The case of blinks, in *Sign Languages: A Cambridge Language Study*, ed. by Diane Brentari, 519–542, Cambridge: Cambridge University Press.
Tateishi, Koichi. (1989) Theoretical implications of Japanese musicians' language, in *Proceedings of the 8th West Coast Conference on Formal Linguistics*, ed. by Jane Fee and Kathryn Hunt, volume 8, 384–398.
Tatham, Mark A. A. (1971) Classifying allophones. *Language and Speech* 14, 140–145.
Taub, Sarah. (2001) *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge: Cambridge University Press.
Tchobanov, Atanas. (2002) *Représentations et apprentissage des primitives phonologiques: Approche neuromiméique*. University of Paris X dissertation.
Templin, M. C. (1957) *Certain Language Skills in Children: Their Development and Interrelationships*. Minneapolis, MN: University of Minnesota Press.
Ten Bosch, L., R. H. Baayen and M. Ernestus. (2006) On speech variation and word type differentiation by articulatory feature representations, in *Proceedings of the 9th International Conference on Spoken Language Processing*, 2230–2233, Pittsburgh, PA.
Tench, P. (1992) *Studies in Systemic Phonology*. London: Pinter Publications.
Teoh, Boon Seong. (1988) *Aspects of Malay Phonology Revisited, a Non-linear Approach.* University of Illinois Urbana-Champaign dissertation.
Tereshchenko, N. M. (1979) *Nganasankii iazyk*. Leningrad: Nauka.
Ternes, Elmar. (1973) *The Phonemic Analysis of Scottish Gaelic*. Forum Phoneticum 1, Hamburg: Helmut Buske Verlag.
Tesar, Bruce. (1995) *Computational Optimality Theory*. University of Colorado dissertation. ROA 90.
Tesar, Bruce. (1996) Computing optimal descriptions for Optimality Theory grammars with context-free position structures, in *34th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference,* 101–107.
Tesar, Bruce. (1998) An iterative strategy for language learning. *Lingua* 104, 131–145.
Tesar, Bruce. (2004) Using inconsistency detection to overcome structural ambiguity. *Linguistic Inquiry* 35, 219–253.
Tesar, Bruce. (2007) Learnability, in *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 555–574, Cambridge: Cambridge University Press.
Tesar, Bruce. (2008) Output-driven maps. ROA 956.
Tesar, Bruce, John Alderete, Graham Horwood, Nazarre Merchant, Koichi Nishitani and Alan Prince. (2003) Surgery in language learning, in *Proceedings of the 22nd Annual Meeting of the West Coast Conference on Formal Linguistics*, ed. by G. Garding and M. Tsujimura, 477–490, Somerville, MA: Cascadilla Press. ROA-619.
Tesar, Bruce and Alan Prince. (2007) Using phonotactics to learn phonological alternations, in *The Panels: Papers from the 39th Annual Meeting of the Chicago Linguistic Society*, ed. by Jon E. Cihlar, Amy L. Franklin, David W. Kaiser, and I. Kimbara, volume 2, 209–237, Chicago, IL: Chicago Linguistic Society.
Tesar, Bruce and Paul Smolensky. (1993) The learnability of Optimality

Theory: An algorithm and some basic complexity results. Technical report, University of Colorado, Boulder. ROA- 2.

Tesar, Bruce and Paul Smolensky. (1998) Learnability in Optimality Theory. *Linguistic Inquiry* 29, 229–268.

Tesar, Bruce and Paul Smolensky. (2000) *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Tesfaye, Ashenafi and Klaus Wedekind. (1990) Characteristics of Omotic tone: Shinasha (Borna). *Studies in African Linguistics* 21, 347–368.

Tessier, Anne-Michelle. (2006a) *Biases and Stages in Phonological Acquisition*. University of Massachusetts, Amherst dissertation.

Tessier, Anne-Michelle. (2006b) Testing for OO-faithfulness in artificial phonological acquisition. Presented at BUCLD 30, Boston University.

't Hart, Johan and René Collier. (1975) Integrating different levels of intonation analysis. *Journal of Phonetics* 1, 309–327.

't Hart, Johan, René Collier, and Antonie Cohen. (1990) *A Perceptual Study of Intonation: An Experimental Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

Thelen, E. (1989) Self-organization in developmental processes: Can systems approaches work?, in *Systems and Development*, ed. by M. R. Gunnar and E. Thelen, 555–591, Hillsdale, NJ: Lawrence Erlbaum.

Thelen, E. (1991) Motor aspects of emergent speech: A dynamic approach, in *Biological and Behavioral Determinants of Language Development*, ed. by N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, and M. Studdert-Kennedy, 339–362, Lawrence Erlbaum.

Thelen, Esther. (1981) Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology* 17, 237–257.

Thomas, Elaine. (1974) Engenni, in *Ten Nigerian Tone Systems*, ed. by John T. Bendor-Samuel, volume 4 of *Studies in Nigerian Languages*, 13–26, Dallas and Arlington, TX: Summer Institute of Linguistics and University of Texas.

Thompson, Robin, Karen Emmorey, and Tamar H. Gollan. (2005) "Tip of the Fingers" Experiences by Deaf Signers: Insights into the Organization of a Sign-Based Lexicon. *Psychological Science* 16, 856–860.

Thompson, Robin, Karen Emmorey, and Tamar H. Gollan. (2006) The relationship between eye gaze and verb agreement in American Sign Language. *Natural Language and Linguistic Theory* 24, 571–604.

Thorn, Annabel S. C. and Clive R. Frankish. (2005) Long-term knowledge effects on serial recall of non-words are not exclusively lexical. *Journal of Experimental Psychology: Learning. Memory, and Cognition* 31, 729–735.

Thorsen, Nina G. (1983) Two issues in the prosody of Standard Danish: The lack of sentence accent and the representation of sentence intonation, in *Prosody: Models and Measurements*, ed. by Anne Cutler and D. Robert Ladd, 27–38, Berlin: Springer.

Thorsen, Nina G. (1984) Intonation and text in Standard Danish – with special reference to the abstract representation of intonation, in *Phonologica 1984: Proceedings of the 5th International Phonology Meeting, Eisenstadt*, ed. by Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer, and John R. Rennison, 301–309, Cambridge: Cambridge University Press.

Thorsen, Nina G. (1985) Intonation and text in Standard Danish. *Journal of the Acoustical Society of America* 77, 1205–1216.

Thorsen, Nina G. (1986) Sentence intonation in textual context – supplementary data. *Journal of the Acoustical Society of America* 80, 1041–1047.

Thurgood, Graham. (2002) Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* 19, 333–363.

Tiersma, P. M. (1982) Local and general markedness. *Language* 58, 832–849.
Tieszen, Bozena. (1997) *Final Stop Devoicing in Polish: An Acoustic and Historical Account for Incomplete Neutralization*. University of Wisconsin, Madison dissertation.
Tifrit, Ali. (2005) *Marque et syllabe: de la representation des fricatives du français dans le modèle contours*. Université de Paris X dissertation.
Tokizaki, Hisao. (2008) *Syntactic Structure and Silence: A Minimalist Theory of Syntax-Phonology Interface*. Tokyo: Hituzi Syobo Publishing.
Topintzi, Nina. (2006) *Moraic Onsets*. University College, London dissertation.
Topintzi, Nina. (2008) On the existence of moraic onsets. *Natural Language and Linguistic Theory* 26, 147–184.
Topping, Donald. (1973) *Chamorro Reference Grammar*. Honolulu, HI: University of Hawaii Press.
Törkenczy, Miklos and Peter Siptar. (1999) Hungarian syllable structure: Arguments for/against complex constituents, in *The Syllable: Views and Facts*, ed. by Harry van der Hulst and Nancy A. Ritter, 249–284, Berlin: Mouton de Gruyter.
van der Torre, Eric-Jan. (2003) *Dutch Sonorants: The Role of Place of Articulation in Phonotactics*. University of Leiden dissertation. Published by LOT, No. 81.
Trager, George L. (1942) The phoneme "t": A study in theory and method. *American Speech* 17, 144–148.
Trager, George L. and Bernard Bloch. (1941) The syllabic phonemes of English. *Language* 17, 223–246.
Traill, A. (1985) *Phonetic and Phonological Studies of !Xóõ Bushman*. Hamburg: Helmut Buske.
Tranel, Bernard. (1991) CVC light syllables, geminates, and moraic theory. *Phonology* 8, 291–302.
Treiman, Rebecca. (1983) The structure of spoken syllables: Evidence from novel word games. *Cognition* 15, 49–74.
Treiman, Rebecca. (1986) The division between onsets and rimes in English syllables. *Journal of Memory and Language* 25, 476–491.
Treiman, Rebecca and Catalina Danis. (1988) Syllabification of intervocalic consonants. *Journal of Memory and Language* 27, 87–104.
Treiman, Rebecca, Brett Kessler, Stephanie Knewasser, Ruth Tincoff, and Margo Bowman. (2000) English speakers' sensitivity to phonotactic patterns, in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, ed. by Michael B. Broe and Janet B. Pierrehumbert, 269–282, Cambridge: Cambridge University Press.
Trentin, Edmondo and Marco Gori. (2001) A survey of hybrid ANN/HMM models for automatic speech recognition. *Neurocomputing* 37, 91–126.
Trommelen, M. (1983) *The Syllable in Dutch, with Special Reference to Diminutive Formation*. Dordrecht: Foris.
Trubetzkoy, Nikolai S. (1939) *Grundzüge der Phonologie*. (Pagination refers to 1958 German edition.) Prague: Travaux du cercle linguistique de Prague.
Truckenbrodt, Hubert. (1995) *Phonological Phrases: Their Relation to Syntax, Focus and Prominence*. MIT dissertation.
Truckenbrodt, Hubert. (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219–256.
Truckenbrodt, Hubert. (2005) A short report on intonation phrase boundaries in German. *Linguistische Berichte* 203, 273–296.
Truckenbrodt, Hubert. (2006) The syntax-phonology interface, in *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 435–456, Cambridge: Cambridge University Press.

Trudgill, P. (1974) *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Tryon, Darrell. (1970) *An Introduction to Maranungku*. Canberra: Australian National University.

Tsereteli, Konstanin. (1982) On one suprasegmental phoneme in Modern Semitic. *Journal of the American Oriental Society* 102, 343–346.

Tsujimura, Natsuko. (2007) *An Introduction to Japanese Linguistics*. Oxford: Wiley-Blackwell, 2nd edition.

Tucker, Archibald N. (1955) *A Maasai Grammar, with Vocabulary*. London: Longmans, Green and Co.

Tuller, Betty and J. A. S. Kelso. (1991) The production and perception of syllabic structure. *Journal of Speech and Hearing Research* 34, 501–508.

Turk, Alice. (1992) The American English flapping rule and the effect of stress on stop consonant durations. *Cornell Working Papers in Phonetics* 7, 103–133.

Tyler, Stephen. (1969) *Koya: An Outline Grammar*. Berkeley, CA: University of California Press.

Uffmann, Christian. (2004) *Vowel Epenthesis in Loanword Phonology*. University of Marburg dissertation.

Ullman, Michael T. (2004) Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92, 231–270.

Ulrich, Charles. (1997) Loanword adaptation in Lama: Testing the TCRS model. *Canadian Journal of Linguistics* 42, 415–463.

Urbanczyk, Suzanne. (1996) *Patterns of Reduplication in Lushootseed*. Amherst, MA: GLSA.

Vago, Robert. (1973) Abstract vowel harmony systems in Uralic and Altaic Languages. *Language* 49, 579–605.

Vago, Robert. (1974) *Hungarian Generative Phonology*. Harvard University dissertation.

Vago, Robert. (1976) Theoretical implications of Hungarian vowel harmony. *Linguistic Inquiry* 7, 242–263.

Vago, Robert. (1985) The treatment of long vowels in word games, in *Phonology Yearbook*, volume 2, 329–342.

Vago, Robert. (1988) *Vowel Harmony in Finnish Word Games*. The Hague: Walter de Gruyter.

Vallabha, G. K., J. L. McClelland, F. Pons, Janet F. Werker, and S. Amano. (2007) Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences* 104, 13273–13278.

Vallar, Giuseppe. (2006) Memory systems: The case of phonological short-term memory. *Cognitive Neuropsychology* 23, 135–155.

Valli, Clayton and Cecil Lucas. (1992) *The Linguistic Structure of American Sign Language*. Washington, DC: Gallaudet University Press.

Van Bael, C., L. Boves, H. van den Heuvel, and H. Strik. (2006) Automatic phonetic transcription of large speech corpora: A comparative study, in *Proceedings of the 9th International Conference on Spoken Language Processing*, 1085–1088, Pittsburgh, PA.

van Gelder, Timothy J. (1998) The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21, 1–14.

Vance, Timothy. (1987) *An Introduction to Japanese Phonology*. Albany, NY: SUNY Press.

Vanderslice, Ralph and Peter Ladefoged. (1972) Binary suprasegmental features and transformational word-accentuation rules. *Language* 48, 819–838.

Vaux, Bert. (1996) Abkhaz Mabkhaz: m-reduplication in Abkhaz and the problem of melodic invariance. Manuscript, Harvard University.

Vaux, Bert. (1998) *The Phonology of Armenian*. Oxford: Clarendon Press.

Vaux, Bert. (2001a) Consonant epenthesis and hypercorrection. Presented at the Annual Meeting of the Linguistic Society of America, Washington, DC.

Vaux, Bert. (2001b) Dispersion theory and the Abkhaz vowel system. Presented at the 32nd Annual Meeting of the North East Linguistic Society.
Vaux, Bert. (2004) [atr] and [back] Harmony in the Altaic languages, in *Investigations into Formal Altaic Linguistics*, ed. by Sergei Tatevosov, 50–67, Moscow: MAKS Press.
Vaux, Bert. (2005) Formal and empirical arguments for morpheme structure constraints. Presented at the Annual Meeting of the Linguistic Society of America, Oakland.
Vaux, Bert. (2006) What second language phenomena reveal about phonological cognition. Presented at NAPhC 4, Montreal.
Vaux, Bert. (2008) Why the phonological component must be serial and rule-based, in *Rules, Constraints, and Phonological Phenomena*, ed. by Bert Vaux and Andrew Nevins, 20–61, Oxford: Oxford University Press.
Vaux, Bert and Benjamin Bruening. (2007) Abkhaz-Mabkhaz: M-Reduplication in Abkhaz. Presented at the Linguistic Society of America Annual Meeting, Chicago.
Vaux, Bert and Scott Golder. (2003) The Harvard Dialect Survey. Questions and maps of results available at http://www.tekstlab.uio.no/cambridge_survey.
Vaux, Bert and Andrew Wolfe. (2009) The appendix, in *Contemporary Views on Architecture and Representations in Phonology*, ed. by Eric Raimy and Charles Cairns, 101–143, Cambridge, MA: MIT Press.
Vaysman, Olga. (2009) *Segmental Alternations and Metrical Theory*. MIT dissertation.
Velten, H. V. (1943) The growth of phonemic and lexical patterns in infant language. *Language* 19, 281–292.
Vendelin, Inga and Sharon Peperkamp. (2006) The influence of orthography on loanword adaptations. *Lingua* 116, 996–1007.
Venditti, Jennifer J. (1997) Japanese ToBI labelling guidelines. *Ohio State Working Papers in Linguistics* 50, 127–62.
Venditti, Jennifer J. (2000) *Discourse Structure and Attentional Salience Effects on Japanese Intonation*. Ohio State University dissertation.
Venditti, Jennifer J. (2005) The J-ToBI Model of Japanese intonation, in *Prosodic Typology: The Phonology of Intonation and Phrasing*, ed. by Sun-Ah Jun, 182–200, Oxford, New York: Oxford University Press.
Venditti, Jennifer J., Sun-Ah Jun, and Mary E. Beckman. (1996) Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English, in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisitions*, ed. by James L. Morgan and Katherine Demuth, 287–311, Mahwah, NJ: Lawrence Erlbaum Associates.
Venditti, Jennifer J., Kazuaki Maeda, and Jan P. H. van Santen. (1998) Modeling Japanese boundary pitch movements for speech synthesis, in *Proceedings of the 3rd ESCA Workshop on Speech Synthesis*, ed. by Mike Edgington, 317–322.
Venditti, Jennifer J., Kikuo Maekawa, and Mary E. Beckman. (2008) Prominence marking in the Japanese intonation system, in *The Oxford Handbook of Japanese Linguistics*, ed. by Shigeru Miyagawa and Mamoru Saito, 456–512, Oxford, New York: Oxford University Press.
Venezky, Richard. (1970) *The Structure of English Orthography*. The Hague: Mouton.
Vennemann, Theo. (1970) The German velar nasal: A case for abstract phonology. *Phonetica* 22, 65–81.
Vennemann, Theo. (1971) The phonology of Gothic vowels. *Language* 47, 90–132.
Vennemann, Theo. (1972a) On the theory of syllabic phonology. *Linguistische Berichte* 18, 1–18.
Vennemann, Theo. (1972b) Phonetic analogy and conceptual analogy, in

Schuchhardt, the Neogrammarians, and the Transformational Theory of Phonological Change: Four Essays by Hugo Schuchhardt, ed. by Theo Vennemann and T. H. Wilbur, 115–179, Frankfurt am Main: Athenaum.

Vennemann, Theo. (1972c) Rule inversion. *Lingua* 29, 209–242.

Vennemann, Theo. (1974) Phonological concreteness in natural generative grammar, in *Toward Tomorrow's Linguistics*, ed. by Roger Shuy and C.-J. N. Bailey, Washington, DC: Georgetown University Press.

Vickers, Douglas and Michael D. Lee. (1998) Dynamic models of simple judgments: I (Properties of a self-regulating accumulator module). *Nonlinear Dynamics, Psychology, and Life Sciences* 2, 169–194.

Vieregge, W. H. (1987) Basic aspects of phonetic segmental transcription, in *Probleme der phonetischen Transkription*, Stuttgart: Franz Steiner Verlag.

Vigário, Marina, Maria Joã Freitas, and Sónia Frota. (2006) Grammar and frequency effects in the acquisition of prosodic words in European Portuguese. *Language and Speech* 49, 175–203.

Vigário, Marina and Sónia Frota. (2002) Prosodic word deletion in coordinate structures. *Journal of Portuguese Linguistics* 1, 241–264.

Vihman, Marilyn M. (1978) Consonant harmony: Its scope and function in child language, in *Universals of Human Language*, ed. by Joseph H. Greenberg, 281–334, Stanford, CA: Stanford University Press.

Vihman, Marilyn M. (1993) Variable paths to early word production. *Journal of Phonetics* 21, 61–82.

Vihman, Marilyn M. (1996) *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

Vihman, Marilyn M., Charles A. Ferguson, and Mary Elbert. (1986) Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics* 7, 3–40.

Viljoen, Johannes. (1973) *Manual for Ndonga, Part 1*. Pretoria: University of South Africa.

Vincent, Nathalie. (2004) *Corpus de l'influence de la graphie dans les noms de marques commerciales anglais utilisés en français québécois*. Quebec City Project CoPho Research Report, Université Laval.

Vitevitch, Michael S. (2003) The influence of sublexical and lexical representations in the processing of spoken words in English. *Clinical Linguistics and Phonetics* 17, 487–499.

Vitevitch, Michael S., Jonna Armbrüster, and Shinying Chu. (2004) Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 514–529.

Vitevitch, Michael S. and Paul A. Luce. (1999) Probabilistic phonotactics and neighborhood density in spoken word recognition. *Journal of Memory and Language* 40, 374–408.

Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer. (1996) Phonotactic and metrical influences on adult ratings of spoken nonsense words, in *Proceedings of the International Conference on Spoken Language Processing*, Philadelphia, PA.

Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer. (1997) Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40, 47–62.

Vitevitch, Michael S., David B. Pisoni, Karen Iler Kirk, Marcia Hay-Cutcheon and Stacey L. Yount. (2002) Effects of phonotactic probabilities on the processing of spoken words and nonwords by adults with cochlear implants who were postlingually deafened. *The Volta Review* 102, 283–302.

Vitevitch, Michael S. and M. Sommers. (2003) The facilitative influence of phonological similarity and neighborhood frequency in speech production. *Memory and Cognition* 31, 491–504.
Voorhoeve, C. L. (1970) The languages of the Lake Murray Area. *Pacific Linguistics A* 25, 1–18.
Voorhoeve, Jan. (1971) Tonology of the Bamileke noun. *Journal of African Languages* 20, 44–53.
Voorhoeve, Jan. (1973) Safwa as restricted tone system. *Studies in African Linguistics* 4, 1–22.
Vousden, Janet I., Gordon D. A. Brown, and Trevor A. Harley. (2000) Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology* 41, 101–175.
Vroomen, Jean, Jurki Tuomainen, and Beatrice de Gelder. (1998) The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language* 38, 133–149.
Wagner, Michael. (2005) *Prosody and Recursion*. MIT dissertation.
Wagner, Michael. (2010) Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory* Volume 28, 183–237.
Waibel, Alexander, Toshiyuki Hanazawa, Geoffrey Hinton, Kiyohiro Shikano, and Kevin J. Lang. (1989) Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37, 1888–1898.
Walker, Douglas C. (1996) The new stability of unstable -e in French. *French Language Studies* 6, 211–229.
Walker, Rachel. (1996) Prominence-driven stress. ROA-172.
Walker, Rachel. (1998) *Nasalization, Neutral Segments, and Opacity Effects*. University of California, Santa Cruz dissertation. Published by Garland Publishing, New York, 2000.
Walker, Rachel. (1999) Guaraní voiceless stops in oral versus nasal contexts: An acoustical study. *Journal of the International Phonetic Association* 29, 63–94.
Walker, Rachel. (2000a) Long-distance consonantal identity effects, in *Proceedings of the 19th Annual Meeting of the West Coast Conference on Formal Linguistics*, ed. by Roger Billery and Brook Lillehaugen, 532–545, Somerville, MA: Cascadilla Press.
Walker, Rachel. (2000b) Yaka nasal harmony: Spreading or segmental correspondence?, in *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society*, ed. by Lisa J. Conathan, Jeff Good, Darya Kavitskaya, Alyssa B. Wulf, and Alan C. L. Yu, 321–332, Berkeley, CA: Berkeley Linguistics Society.
Walker, Rachel. (2001a) Positional markedness in vowel harmony, in *Proceedings of HILP 5*, ed. by C. Féry, A. Dubach Green, and R. van de Vijver, 212–232, University of Potsdam.
Walker, Rachel. (2001b) Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language and Linguistic Theory* 19, 827–878.
Walker, Rachel. (2004) Vowel feature licensing at a distance: Evidence from northern Spanish language varieties, in *Proceedings of the 23rd Annual Meeting of the West Coast Conference on Formal Linguistics*, 787–800.
Walker, Rachel. (2005) Weak triggers in vowel harmony. *Natural Language and Linguistic Theory* 23, 917–989.
Walker, Rachel. (2007) Nasal and oral consonant similarity in speech errors: Exploring parallels with nasal consonant harmony. *Language and Cognitive Processes* 22, 1–41.
Walker, Rachel, Dani Byrd, and Fidèle Mpiranya. (2008) An articulatory view of Kinyarwanda coronal harmony. *Phonology* 25, 499–535.
Walker, Rachel and Fidèle Mpiranya. (2006) On triggers and opacity in

coronal harmony, in *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*, 383–394.
Walker, Rachel and Geoffrey K. Pullum. (1999) Possible and impossible segments. *Language* 75, 764–780.
Walsh, Margaret A. and Randy L. Diehl. (1991) Formant transition duration and amplitude rise time as cues to the stop/glide distinction. *The Quarterly Journal of Experimental Psychology* 43A, 603–620.
Walter, Mary Ann. (2003) Kalaam, Kalaarbaam: An Arabic language game in Hadramaut, Yemen, in *Proceedings of the 10th Annual Symposium about Language and Society – Austin*, volume 45, Austin, TX: Texas Linguistic Forum.
Walter, Mary Ann. (2007) Boudledidge: A contribution to language game typology. Poster at the 15th Manchester Phonology Meeting, Manchester, May 24–26.
Wan, I-Ping and Jeri Jaeger. (1998) Speech errors and the representation of tone in Mandarin Chinese. *Phonology* 15, 417–461.
Wang, William S.-Y. (1969) Competing changes as a cause of residue. *Language* 45, 9–25.
Wang, William S.-Y. and Charles Fillmore. (1961) Intrinsic cues and consonant perception. *Journal of Speech and Hearing Research* 4, 130–136.
Wanlass, Zara. (2008) Evidence from online corpora for variation in Hungarian vowel harmony, in *Approaches to Hungarian 10*, ed. by Christopher Piñon, Péter Siptar, and Szilárd Szentgyörgyi, Szeged: JATE Press.
Ward, Ida C. (1948) Verbal tone patterns in West African languages. *Bulletin of the School of Oriental and African Studies, University of London* 12, 831–837.
Warker, Jill and Gary Dell. (2006) Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32, 387–398.
Warner, Natasha, Erin Good, Allard Jongman, and Joan Sereno. (2006) Letter to the editor: Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics* 34, 285–293.
Warner, Natasha, Allard Jongman, Joan Sereno, and Rachel Kemps. (2004) Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics* 32, 251–276.
Wasow, Tom. (1978) On constraining the class of transformational languages. *Synthese* 39, 81–104.
Waterson, Natalie (1971) Child phonology: A prosodic view. *Journal of Linguistics* 7, 179–211.
Waterson, Natalie (1987) *Prosodic Phonology: The Theory and its Application to Language Acquisition and Speech Processing*. Tübingen: Gunter Narr Verlag.
Watson, Janet. (1999) The directionality of emphasis spread in Arabic. *Linguistic Inquiry* 30, 289–300.
Watson, Janet. (2002) *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.
Watson, John B. (1924) *Behaviorism*. New York: Norton.
Watters, James. (1988) *Topics in Tepehua Grammar*. University of California, Berkeley dissertation.
Webb, Charlotte. (1974) *Metathesis*. University of Texas, Austin dissertation.
Weber, David and Wesley Thiesen. (2000) A synopsis of Bora tone. Manuscript SIL.
Wedel, Andrew. (2004) *Self-organization and Categorical Behavior in Phonology*. University of California, Santa Cruz dissertation.
Wedel, Andrew. (2006) Exemplar models, evolution, and language change. *The Linguistic Review* 23, 247–274.
Wedel, Andrew. (2007) Feedback and regularity in the lexicon. *Phonology* 24, 147–185.

Wedel, Andrew and Adam Ussishkin. (2002) Neighborhood density and the root-affix distinction, in *Proceedings of the 32nd Annual Meeting of the North East Linguistic Society*, ed. by Masako Hirotani, 539–549.

Weijer, Jeroen Maarten van de. (1996) *Segmental Structure and Complex Segments*. Tübingen: Niemeyer.

Weinreich, Uriel, William Labov, and Marvin Herzog. (1968) Empirical foundations for a theory of language change, in *Directions for Historical Linguistics. A Symposium*, ed. by W. Lehmann and Y. Malkiel, 95–188, Austin, TX: University of Texas Press.

Weissenborn, Jürgen and Barbara Höhle. (2001) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, Amsterdam, Philadelphia: Benjamins.

Welch, Robert B. and David H. Warren. (1986) Intersensory interactions, in *Handbook of Perception and Human Performance, Volume 1: Sensory Processes and Perception*, ed. by L. Kaufman, 25–36, New York: John Wiley & Sons.

Wells, Rulon. (1947) Immediate constituents. *Language* 23, 81–117.

Welmers, William E. (1959) Tonemics, morphotonemics, and tonal morphemes. *General Linguistics* 4, 1–9.

Welmers, William E. (1962) The phonology of Kpelle. *Journal of African Languages* 1, 69–93.

Welmers, William E. (1973) *African Language Structures*. Berkeley and Los Angeles: University of California Press.

Werle, Adam. (2009) *Word, Phrase and Clitic Prosody in Bosnian, Serbian and Croatian*. University of Massachusetts, Amherst dissertation.

Wester, Mirjam, Judith M. Kessens, Catia Cucchiarini, and Helmer Strik. (2001) Obtaining phonetic transcriptions: A comparison between expert listeners and a continuous speech recognizer. *Language and Speech* 44, 377–403.

Wetzels, W. Leo and Joan Mascaró. (2001) The typology of voicing and devoicing. *Language* 77, 207–244.

Wheeler, Deirdre W., and David S. Touretzky. (1997) A parallel licensing model of normal slips and phonemic paraphasias. *Brain and Language* 59, 147–201.

White, Katherine S. and James L. Morgan. (2008) Subsegmental detail in early lexical representations. *Journal of Memory and Language* 59, 114–132.

White, Robin. (1973) *Klamath Phonology*. Seattle, WA: Department of Linguistics, University of Washington.

Whitney, William Dwight. (1874) *Oriental and Linguistic Studies*. New York: Scribner, Armstrong.

Whitney, William Dwight. (1889) *Sanskrit Grammar*. London: Oxford University Press.

Whorf, Benjamin Lee. (1940) Linguistics as an exact science. *Technology Review* 41, 80–83.

Wiebe, Bruce. (1992) Modelling autosegmental phonology with multi-tape finite state transducers. Master's thesis, Simon Fraser University. ftp://ftp.fas.ca/pub/cs/theses/1992/Bruce/WiebeMSc.ps.gz

Wiese, Richard. (1996) *The Phonology of German*. Oxford: Oxford University Press.

Wilbur, Ronnie B. (1973) *The Phonology of Reduplication*. University of Illinois, Urbana-Champaign dissertation.

Wilbur, Ronnie B. (1994) Eye blinks and ASL phrase structure. *Sign Language Studies* 84, 221–240.

Wilbur, Ronnie B. (1999) Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech* 42, 229–250.

Wilbur, Ronnie B. (2008) Complex predicates involving events, time, and aspect: Is this why sign languages look so similar?, in *Signs of the Time: Selected Papers from TISLR 2004*, ed. by Josep Quer, 219–250, Hamburg: Signum Verlag.

Wilbur, Ronnie B. (2010) The event visibility hypothesis, in *Sign Languages: A Cambridge Language Survey*, ed. by Diane Brentari, 357–382, Cambridge: Cambridge University Press.

Wilbur, Ronnie B., Edward Klima, and Ursula Bellugi. (1983) Roots: On the search for the origins of signs in ASL, in *Proceedings from the 38th Annual Meeting of the Chicago Linguistic Society*, 314–336, Chicago: Chicago Linguistic Society.

Wilbur, Ronnie B. and Susan Nolen. (1986) The duration of syllables in American Sign Language. *Language and Speech* 29, 263–280.

Wilbur, Ronnie B. and Cynthia Patschke. (1998) Body leans and marking contrast in ASL. *Journal of Pragmatics* 30, 275–303.

Wilcox, Phyllis. (2001) *Metaphor in American Sign Language*. Washington, DC: Gallaudet University Press.

Wilcox, Sherman, Paolo Rossini, and Elena A. Pizzuto. (2010) Grammaticalization in sign languages, in *Sign Languages: A Cambridge Language Survey*, ed. by Diane Brentari, 332–356, Cambridge, UK: Cambridge University Press.

Wilkinson, Karina. (1988) Prosodic structure and Lardil phonology. *Linguistic Inquiry* 19, 325–334.

Willard, Rainbow. (2004) Dominance effects in a dialect of Mam Maya. Manuscript, University of California Berkeley.

Willerman, Raquel. (1994) *The Phonetics of Pronouns: Articulatory Bases of Markedness*. University of Texas, Austin dissertation.

Williams, Lee. (1979) The modification of speech perception and production in second-language learning. *Perception and Psychophysics* 26, 95–104.

Williamson, Kay. (1986) The Igbo associative and specific constructions, in *The Phonological Representation of Suprasegmentals*, ed. by Koen Bogers, Harry van der Hulst, and Maarten Mous, 195–208, Dordrecht: Foris.

Williamson, Kay. (1988) Tone and accent in Ijo, in *Autosegmental Studies on Pitch Accent*, ed. by Harry van der Hulst and Norval Smith, 253–278, Dordrecht: Foris.

Wilson, Colin. (2002) Unbounded spreading is local spreading iterated unboundedly. Talk presented at the University of Southern California.

Wilson, Colin. (2003) Experimental investigation of phonological naturalness, in *Proceedings of the 22nd Annual Meeting of the West Coast Conference on Formal Linguistics*, 533–546, Somerville, MA: Cascadilla Press.

Wilson, Colin. (2006a) Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30, 945–982.

Wilson, Colin. (2006b) The Luce Choice Ranker. Manuscript, Department of Linguistics, UCLA.

Wilson, Colin. (2006c) Unbounded spreading is myopic. Presented at *Current Perspectives on Phonology Workshop*, Indiana University.

Wingfield, Arthur. (1978) *Human Memory and Learning: An Introduction*. New York: Harper and Row Publishers.

Winitz, Harris, M. E. Scheib, and James A. Reeds. (1972) Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *The Journal of the Acoustical Society of America* 51, 1309–1317.

Winograd, Terry. (1983) *Language as a Cognitive process, Volume I: Syntax*. Reading, MA: Addison-Wesley.

Withgott, Margaret M. and Francine R. Chen. (1993) *Computational Models of American Speech*. Stanford, CA: CSLI.

Wolf, Matt. (2008) *Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-based Model*. University of Massachusetts, Amherst dissertation.

Wong, Wai-Yi Peggy. (2006) *Syllable Fusion in Hong Kong Cantonese Connected Speech*. Ohio State University dissertation.

Wong, Wai-Yi Peggy, Marjorie K.-M. Chan, and Mary E. Beckman. (2005) An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese, in *Prosodic Typology: The Phonology of Intonation and Phrasing*, ed. by Sun-Ah Jun, 271–300, Oxford: Oxford University Press.

Woodbury, Anthony. (1987) Meaningful phonological processes: A consideration of Central Alaskan Yupik Eskimo prosody. *Language* 63, 685–740.

Wright, Richard. (2004) Factors of lexical competition in vowel articulation, in *Papers in Laboratory Phonology VI: Phonetic Interpretation*, ed. by John Local, Richard Ogden, and Rosalind Temple, 26–50, Cambridge: Cambridge University Press.

Wurzel, Wolfgang U. (1987) Paradigmenstrukturbedingungen: Aufbau und Veranderung von Flexionsparadigmen, in *Papers from the 7th International Conference on Historical Linguistics*, ed. by A. G. Ramat, O. Carruba, and G. Bernini, 629–644, Amsterdam: Benjamins.

Xu, Yi and Fang Liu. (2007) Determining the temporal interval of segments with the help of F0 contours. *Journal of Phonetics* 35, 398–420.

Xu, Yi and Emily Wang. (2001) Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication* 33, 319–337.

Xu, Yisheng, Krishnan Ananthanarayan, and Jackson T. Gandour. (2006) Specificity of experience dependent pitch representation in the brainstem. *Neuroreport* 17, 1601–1605.

Yamada, Reiko and Y. Tohkura. (1992) The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception and Psychophysics* 52, 376–392.

Yasinnik, Yelena, Margaret Renwick, and Stefanie Shattuck-Hufnagel. (2004) The timing of speech-accompanying gestures with respect to prosody. Presented at the *From Sound to Sense Conference*.

Yau, Shun-chiu. (1980) Sentential connotations in Cantonese. *Fangyan* 80, 35–52.

Yavas, M. S. (1994) *First and Second Language Phonology*. San Diego, CA: Singular Publishing Group.

Yeni-Komshian, G. H., J. F. Kavanagh, and C. A. Ferguson. (1980) *Child Phonology*, volume 1. New York: Academic Press.

Yi, Kwangoh. (1999) The internal structure of Korean syllables: Rhyme or body? *Korean Journal of Experimental and Cognitive Psychology* 10, 67–83.

Yip, Moira. (1980) *The Tonal Phonology of Chinese*. MIT dissertation. Published by Garland Publishing, New York, 1990.

Yip, Moira. (1982) Reduplication and C-V skeleta in Chinese secret languages. *Linguistic Inquiry* 13, 637–661.

Yip, Moira. (1988) The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19, 65–100.

Yip, Moira. (1989) Contour tones. *Phonology* 6, 149–174.

Yip, Moira. (1992) Reduplication with fixed melodic material, in *Proceedings of the 22nd Annual Meeting of the North East Linguistic Society*, 459–476, Amherst, MA: GLSA.

Yip, Moira. (1993) Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2, 261–291.

Yip, Moira. (1995) Tone in East Asian languages, in *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 476–494, Oxford: Blackwell.

Yip, Moira. (1996) Lexical optimization in languages without alternations, in *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand and Bernard Laks, 354–385, Salford: ESRI.

Yip, Moira. (1997) Repetition and its avoidance: The case of Javanese, in *Arizona Phonology Conference Proceedings, Vol. 5: Proceedings of the South Western Workshop on Optimality Theory*, ed. by K. Suzuki and D. Elzinga, 238–262,

Tucson, AZ: University of Arizona, Department of Linguistics Coyote Papers.

Yip, Moira. (1998) Identity avoidance in phonology and morphology, in *Morphology and its Relations to Phonology and Syntax*, ed. by Steve LaPointe, Diane Brentari, and Patrick Farrell, 216–246, Stanford, CA: CSLI.

Yip, Moira. (2002) *Tone*. Cambridge Textbooks in Linguistics, Cambridge: Cambridge University Press.

Yip, Moira. (2003) Casting doubt on the Onset/Rime distinction. *Lingua*, 113, 779–816.

Yoshida, Shohei. (1990) *Some Aspects of Governing Relations in Japanese Phonology*. School of Oriental and African Studies, University of London dissertation.

Yoshida, Shohei. (1993) Licensing of empty nuclei: The case of Palestinian vowel harmony. *The Linguistic Review* 10, 127–159.

Younes, Munther. (1991) Emphasis spread in three Arabic dialects. Manuscript, Cornell University.

Younes, Munther. (1993) Emphasis spread in two Arabic dialects, in *Perspectives on Arabic Linguistics V*, ed. by Mushira Eid and Clive Holes, 119–145, Amsterdam: Benjamins.

Yu, Alan C. L. (2000) Stress assignment in Tohono O'odham. *Phonology* 17, 117–135.

Yu, Alan C. L. (2004a) Explaining final obstruent voicing in Lezgian: Phonetics and history. *Language* 80, 73–97.

Yu, Alan C. L. (2004b) Reduplication in English Homeric infixation, in *Proceedings of the 34th Annual Meeting of the North East Linguistic Society*, ed. by Keir Moulton and Matthew Wolf, 619–633, Amherst, MA: GLSA.

Yu, Alan C. L. (2005) Quantity, stress and reduplication in Washo. *Phonology* 22, 437–475.

Yu, Alan C. L. (2007a) *A Natural History of Infixation*. Oxford: Oxford University Press.

Yu, Alan C. L. (2007b) Understanding near-mergers: The case of morphological tone in Cantonese. *Phonology* 24, 187–214.

Yu, Alan C. L. (2007c) Tonal phonetic analogy, in *Proceedings of the International Congress of the Phonetic Sciences XVI*, ed. by Jürgen Trouvain and William J. Barry, 1749–1752.

Yu, Alan C. L. (2008) On iterative infixation, in *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. by Hannah J. Haynie and Charles B. Chang, 516–524, Somerville, MA: Cascadilla Press.

Zaba, Aleksandra. (2008) *Relative Frequency of Patterns and Learnability: The Case of Phonological Harmony*. University of Utah dissertation.

Zamuner, Tania S., LouAnn Gerken, and Michael Hammond. (2004) Phonotactic probabilities in young children's speech production. *Journal of Child Language* 31, 515–536.

Zamuner, Tania S., Annemari Kerkhoff, and Paul Fikkert. (2006) Acquisition of voicing neutralization and alternations in Dutch, in *Proceedings of the 30th Boston University Conference on Language Development*, ed. by David Bamman, Tatiana Magnitskaia, and Colleen Zaller, 701–712, Somerville, MA: Cascadilla Press.

Zawaydeh, Bushra A. (1999) *The Phonetics and Phonology of Gutturals in Arabic*. Indiana University dissertation.

Zazzo, René. (1960) *Les jumeaux: Le couple et la personne*. Paris: Presses Universitaires de France.

Zec, Draga. (1988) *Sonority Constraints on Prosodic Structure*. Stanford University dissertation.

Zec, Draga. (1995a) The role of moraic structure in the distribution of segments within syllables, in *Frontiers of Phonology: Atoms, Structures, Derivations*, ed. by Jacques Durand and Francis Katamba, 149–179, Harlow, UK: Longman.

Zec, Draga. (1995b) Sonority constraints on syllable structure. *Phonology* 12, 85–129.

Zec, Draga. (1999) Footed tones and tonal feet: Rhythmic constituency in a pitch accent language. *Phonology* 16, 225–264.
Zec, Draga. (2003) Prosodic weight, in *The Syllable in Optimality Theory*, ed. by Caroline Féry and Ruben van de Vijver, 123–143, Cambridge: Cambridge University Press.
Zee, Eric. (1990) Vowel devoicing in Shanghai. *Chinese University of Hong Kong Papers in Linguistics* 2, 69–104.
Zee, Eric and Ian Maddieson. (1980) Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. *Glossa* 14, 45–88.
Zepeda, Ofelia. (1983) *A Papago Grammar*. Tucson, AZ: University of Arizona Press.
Zepeda, Ofelia. (1984) *Topics in Papago Morphology*. University of Arizona dissertation.
Zerbian, Sabine. (2006) *Expression of Information Structure in the Bantu Language Northern Sotho*. Humboldt University dissertation.
Zerbian, Sabine. (2007) Phonology phrasing in Northern Sotho. *The Linguistic Review* 24, 233–262.
Zhang, Jie. (2001) *The Effects of Duration and Sonority on Contour Tone Distribution: Typological Survey and Formal Analysis*. Los Angeles: UCLA dissertation.
Zhang, Jie. (2002) *The Effects of Duration and Sonority on Contour Tone Distribution – A Typological Survey and Formal Analysis*. New York: Routledge.
Zhang, Jie, Yuwen Lai, and Craig Turnbull-Sailor. (2006) Wug-testing the "tone circle" in Taiwanese, in *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, and Michael Scanlon, 453–461, Somerville, MA: Cascadilla Press.
Zhang, Xi. (1996) *Vowel Systems of the Manchu-Tungus Languages of China*. University of Toronto dissertation.
Zhukova, Alevtina Nikodimovna. (1972) *Grammatika Koriakskogo Iazyka: Fonetika, Morfologiia*. Leningrad: Izdatelstvo Nauka.
Zimmer, Karl E. (1969) Psychological correlates of some Turkish morpheme structure conditions. *Language* 45, 309–321.
Zinc, Gaston. (1986) *Phonétique historique du français*. Paris: Presses Universitaires de France.
Zoll, Cheryl. (1996) *Parsing Below the Segment in a Constraint Based Framework*. University of California, Berkeley dissertation.
Zoll, Cheryl. (1998) Positional asymmetries and licensing. MIT, dissertation.
Zonneveld, Wim. (1978) *A Formal Theory of Exceptions in Generative Phonology*. Dordrecht: Foris.
Zonneveld, Wim. (1984) The game of the name: Expletive insertion in English. *Linguistic Analysis* 13, 55–60.
Zonneveld, Wim. (2007) Issues in Dutch devoicing: Positional faithfulness, positional markedness, and local conjunction, in *Voicing in Dutch*, ed. by Erik-Jan van der Torre and Jeroen van de Weijer, 1–40, Amsterdam: Benjamins.
Zsiga, Elizabeth C. (1995) An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English, in *Papers in Laboratory Phonology IV*, 282–302, Cambridge: Cambridge University Press.
Zsiga, Elizabeth C. (1997) Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface. *Language* 73, 227–274.
Zsiga, Elizabeth C., Maria Gouskova, and Maria Tlale. (2006) On the status of voiced obstruents in Tswana: Against *ND, in *Proceedings of the 36th Annual Meeting of the North East Linguistic Society*, ed. by Christopher Davis, Amy Rose Deal, and Youri Zabbal, 721–734, Amherst, MA: GLSA.
Zubizarreta, Maria Luisa. (1979) Vowel harmony in Andalusian Spanish, in *MIT Working Papers in Linguistics*, ed. by Ken Safir, volume 1, 1–11, MIT Press.

Zubizarreta, Maria Luisa. (1998) *Prosody, Focus and Word Order*. Cambridge, MA: MIT Press.

Zue, Victor and M. Lafferriere. (1979) Acoustical study of medial /t/, /d/ in American English. *Journal of the Acoustical Society of America*, 66, 1039–1050.

Zuraw, Kie. (2000) *Patterned Exceptions in Phonology*. UCLA dissertation.

Zuraw, Kie. (2002) Aggressive reduplication. *Phonology* 19, 395–439.

Zuraw, Kie. (2007) The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83, 277–316.

Zwicky, Arnold M. (1965) *Topics in Sanskrit Phonology*. MIT dissertation.

Zwicky, Arnold M. (1970a) Class complements in phonology. *Linguistic Inquiry* 1, 262–265.

Zwicky, Arnold M. (1970b) The Free-Ride Principle and two rules of complete assimilation in English, in *Papers from the 6th Regional Meeting of the Chicago Linguistic Society*, 579–588, Chicago: Chicago Linguistic Society.

Zwicky, Arnold M. (1981) "Internal" and "external" evidence in linguistics, in *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 598–604.

Zwicky, Arnold M. (1985a) Heads. *Journal of Linguistics* 21, 1–29.

Zwicky, Arnold M. (1985b) How to describe inflection, in *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, ed. by M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Feder, 372–386, Berkeley, CA: Berkeley Linguistics Society.

Zwicky, Arnold M. and Geoffrey Pullum. (1987) Plain and expressive morphology, in *Berkeley Linguistics Society: Proceedings of the 13th Annual Meeting, General Session and Parasession on Grammar and Cognition*, ed. by Jon Aske, Natasha Beery, Laura Michaelis, and Hana Filip, 330–340, Berkeley, CA: Berkeley Linguistics Society.

Zwirner, Eberhard and Kurt Zwirner. (1966) *Grundfragen der Phonometrie*. Basel: S. Karger.

Index

References to notes and figures are entered as (respectively) 21n or 21f.

α-α relations, 547
∀ (universal quantifier), 39n
∃ (existential quantifier), 39n
(ι) *see* Intonational Phrase (ι)
(φ) *see* Phonological Phrase (φ)
(ω) *see* Prosodic Word (ω)

A-over-A Principle, 6
AAVE (African American Vernacular English), 422f
Abad, May, 408
Abasheikh, Mohamma Imam, 458
Abbi, Anvita, 95
Abboud, Peter, 275
abbreviatory conventions, 11–12
Abkhaz, 96, 731
ablaut, 78–80, 109–10
"absolute slicing" hypothesis, 367
abstractionist models, 388–9
abstractness controversy, 370–1
accent, definition, 532n
Accentual Phrase (AP), 517–18, 528–30, 532n
accommodation, 323, 327–8
acoustic cues
  in language acquisition, 591–2
  misperception of, 324, 326–7, 328–31, 332, 343, 347n
  storage of, 395–7
acquisition, 8–9, 571–95
  babbling, 576–7, 695–6
  bilingual acquisition, 572, 594
  chain shifts, 369–70, 573
  child-directed talk, 499, 584–5, 588, 701
  cognitive and biological models, 576–8
  conference papers, 578–9
  constraints, 574, 576, 578–80, 583–7, 621, 665–72, 685–8, 689n–90n, 737–8
  and covert contrast, 364, 575–6, 592
  diary studies, 572, 574
  first words, 576–7, 579–80
  frequency effects, 580–1, 583, 584–5, 588–90
  learnability, 40–1, 54–7, 64–5, 594–5, 661–88, 689n–90n, 723–36
    nativist theory, 573, 663

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu

© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.

learner simulations, 316, 396–7, 398, 399–400, 415–27, 434n, 594–5, 622–5, 663, 670–1, 685–8
learning bias, 342–5, 688, 729
morphology, 585–7, 675–85
and opacity, 65n, 369–70
phonological alternations, 672–85
phonotactics, 663–72
prosodic features, 574–5, 583–7, 592
second language acquisition, 56, 369–70, 488
sign languages, 695–6, 701, 709, 711–12
and sound change, 322–6, 329, 334–5, 338–9, 347n
stress systems, 162–3, 339, 347n–8n
universals, 573, 577, 666, 685–8, 723–4
variation, 576–7, 582, 591, 662–3, 723–4, 725–34
Adaptive Resonance Theory, 398, 399
adequacy, definition, 37n *see also* evaluation
Adjacency-Identity Constraint (AIC), 14
aerodynamics, 324, 488, 494–5
aesthetic adequacy, definition, 37n
affixes, 68–73, 76–90, 100–1, 569n
dominant vs. recessive, 69–71, 72–3
mobile affixes, 87
root-faithfulness principle, 82
in sign languages, 702–5, 712–17
and vowel harmony, 215–16, 217, 218–19, 255, 280–1, 284–7
African American Vernacular English (AAVE), 422f
Africanist transcription conventions (tone transcription), 499
agent-based simulations, 333, 336
agentives, 85, 89–90, 284
agglutinative languages, layering effects, 75
Aghem, 211–12, 220
agreement
in sign languages, 712–17, 721n
tone agreement, 229
Ahn, Mee-Jin, 133
Aissen, Judith, 175
Ajíbóyè, Oladiipo, 247
Akan, 52
Akinlabi, Akinbiyi, 96
Aksu-Koç, Ahyan, 673
Al-Ani, S. H., 247
Al-Masri, M., 247, 248
Al-Mozainy, Hamza, 726
Al-Sayyid Bedouin Sign Language (ABSL), 711, 712
Alario, François-Xavier, 653, 654
Albert, Salich Y., 93
Albright, Adam, 344, 391, 662, 681, 682, 683, 685, 687
Albro, Daniel M., 666
Alderete, John, 72, 73, 82, 95, 96, 677
algorithms, 415–16, 419–20, 621, 630n, 634, 635, 665–72, 681–2, 685–8, 689n
Align XP theories, 458–68, 483n
alignment constraints, 264–6, 450 *see also* edge constraints
Allen, George D., 574–5
Allen, John, 600
allophones, 352
acquisition, 672–85
affixes, 68–73, 76–90
modeling of, 29–30, 72–3, 76–83, 307–16, 361, 363–4, 662–3
alpha-alpha relations (α-α relations), 547
alphabets, 304, 372n, 499, 504, 766 *see also* notation
Alphei, Henning, 704
Alternation Condition, 80–1
Altschuler, Daniel, 142–3
alveolar consonants, 242, 300
ambiguity, 605–11, 615, 648–9, 729–30 *see also* variation
ambisyllabicity, 196n, 548, 740
American English
corpora, 375–6, 377, 378, 380–1, 387
flaps, 191–2, 306, 312–13, 394, 395
glides, 754
intonation, 376
language games, 737, 744, 745–6
reduction, 378
syllable structure, 165
t/d deletion, 405, 406–7, 409–13, 421–3
vowels, 305, 361, 362f
American Sign Language (ASL), 696, 700, 709, 721n
Amo, 222
Analogical Model, 390, 391
analysis-synthesis method, 495–7, 510, 518, 599–600

analyst bias, 313, 360, 372n, 382–4
Anceaux, Johannes Cornells, 299
anchoring
  language games, 729–30, 739
  tones, 500–1, 506–15, 525
Andersen, Torben, 242
Anderson, Anne H., 380
Anderson, Gregory, 251, 286
Anderson, John M., 534, 543, 548, 566, 567n, 602, 631
Anderson, Stephen R., 320, 321, 341, 453, 744
anti-faithfulness constraints, 72
Anttila, Arto, 57, 63, 67n, 72, 73, 99, 405, 407, 408, 409, 413, 425, 434n
Aoyama, Katsura, 488
aperture changes (sign language), 695
aphasia, 653–4
apocope, 27, 28, 45
Apoussidou, Diana, 675, 677
appendices, 188–90, 552, 555
  definition, 188
Applecross Scottish Gaelic, 246, 270–1
Applegate, Richard, 101, 242
Arabela, 118, 136n
Arabic, 15, 53, 79, 608–9, 726
  Abu Shusha Palestinian Arabic, 248
  Bedouin Arabic, 275, 726
  Cairene Arabic, 113, 140n, 247, 248
  Jordanian Arabic, 247, 248, 271–2
  language games, 741
  loanword adaptation, 760, 765, 766, 773, 775–6
  Meccan Arabic, 53, 113
  Moroccan Arabic, 726, 760, 765, 766, 775–6
  Muscat Arabic, 15
  Najdi Arabic, 275
  Palestinian Arabic, 248, 281
  prosodic features, 91
  San'ani Arabic, 138n
  syllable structure, 112–13, 189, 619–20
  vowel harmony, 247, 248, 275, 281
  Yemeni Arabic, 247
Aramaic, 247–8, 290n
Arbisi-Kelm, Timothy, 487
Archangeli, Diana, 24–5, 53, 92, 216, 247, 254, 255, 259, 260, 265–6, 270, 280, 281–2, 285, 320, 541, 587
archiphonemes, definition, 295–6
Arigibi Kiwai, 232–3
Armbrüster, Jonna, 653
Armenian, 83, 84, 731, 749n
  Western Armenian, 83, 84
Armstrong, Mike, 679
Aronoff, Mark, 101, 196n, 683, 711
Arrernte, 117–18
Arsenault, Paul, 243
articles, 586, 592
articulation
  and aerodynamics, 324, 488
  and contrast reduction, 300, 312, 328
  and covert contrast, 364
  and harmony systems, 254–5, 261, 263, 266, 273, 288, 290n
  in language acquisition, 577, 591–2
  modeling of, 327, 332, 395–7, 434n, 503, 504–5, 516, 660n
  and phonological alternations, 174, 194, 405
  in sign languages, 691–2, 693–701, 720n, 721n
  and sound change, 323–5, 326–31, 332
  subglottal pressure, 505
Articulatory Phonology, 263, 369, 510
artificial language-learning tasks, 343–4, 391, 728–9, 733, 737–8, 746–7
Arvaniti, Amalia, 513–14, 527
Askins, D., 713, 714
aspect, marking of, 203
aspiration, 190, 244, 307, 359
Assamese, 272, 281
assibilation, 14–16, 80
assimilation, 61–2, 327, 373n
  coronal consonants, 11, 295
  corpus studies, 375–6
  feeding/bleeding relationships, 49–50, 66n
  nasals, 9–10, 12, 174, 295, 323, 326–7, 404–5, 407, 488
  and point of articulation, 174
  sign languages, 698
  tones, 208, 211–13, 214–17, 220–30
  see also harmony systems
associative morpheme (genitive case marker), 204, 210
Asturian Lena, 254
Asudeh, A., 621
Atomic Phonology, 23–4

auditory processing, 704–5
Auer, Peter, 172
Auger, Julie, 139n
augmentation processes (acquisition), 574–5
Automatic Speech Recognition systems, 380, 383–4, 599–600, 610–11, 625, 627, 630n
Autosegmental-Metrical Phonology, 456, 487–91, 500, 507, 513, 515–21
Autosegmental Phonology, 18–22, 24, 297, 502, 510
autosegmental representation
  definition, 367
  finite state automata, 608–11, 614–15
  harmony systems, 256–61, 264–5, 277–8
  and language acquisition, 574, 579
  quantity, 104–8
  tones, 206–13, 238, 366–7, 483n–4n, 502–21
Aveledo, Fraibet, 673
avoidance, 112–16, 120–2, 730–4
Awad Bing, 232
Azerbaijani Jewish Aramaic, 247–8, 290n
Azuma, Tamiko, 627

Baayen, R. Harald, 308, 386, 390, 427, 430, 681
babbling, 576–7, 695–6
Bach, Emmon, 170, 336
backcopying, 101–2, 729
backness (vowels), 251–2, 264–5, 268–9, 290n
Baertsch, Karen, 135n
Bagemihl, Bruce, 557, 722, 730, 735, 742–4, 749n
Bailey, Todd M., 641, 669
Baker, Brett, 111–12, 272, 293
Baković, Eric, 53, 60, 61–2, 66n, 152, 277, 279, 280
Bakwiri, 735
Balling, Laura, 386
Balog, Heather, 701
Baltaxe, Christiane, 352, 371n, 372n
Bamileke-Dschang, 210
Banfield, A. W., 200
Bantu languages, 440–50
  acquisition, 579, 585
  harmony systems, 243, 244, 249, 253, 273, 285
  tones, 211–12, 215–17, 223–4, 233–4
Bao, Zhiming, 209
Barasana, 229
Bard, Ellen Gurman, 640
Barker, M. A. R., 153–4
Barlow, Jessica, 56, 369, 581, 730
Barnes, Janet, 246
Barnes, Jonathan, 327, 336
Barnes, Jonathan Allen, 294, 302, 431
Baroni, Marco, 314–15
Barton, David, 364, 575
Basbøll, Hans, 135n
Base-Identity Constraints, 76
Base-Reduplicant Correspondence Theory, 95, 101–2
Basque, 46, 473–8
  Lekeitio Basque, 473–8, 484n
  Western Basque, 46
Baudouin de Courtenay, Jan, 321
Bauer, Laurie, 284
Bayley, Robert, 407, 424
Beaulieu, Kathy, 768
Becker, Michael, 685–6
Beckman, Jill, 265, 273, 283
Beckman, Mary E., 320, 365, 456, 460, 487, 507, 509, 510, 513, 515, 516, 518, 520, 523, 528, 532n
Beddor, Patrice S., 323
Bedouin Sign Language (Al-Sayyid Bedouin Sign Language), 711, 712
behavioral experiments, 636–45
behaviorism, 572
Béland, Renée, 776
Belarusian (Belorussian), 294, 302–4
Belhare (Belhariya), 69
Bell, Alan, 377
Bell, C. G., 599
Bella Coola, 557
Bellugi, Ursula, 709, 711–12
Belorussian (Belarusian), 294, 302–4
Bendor-Samuel, J. T., 293
Benedicto, Elena, 716–17
Bengali, 282, 467
Benus, Stefan, 288, 335
Berber, 176–7, 504–5, 557
Berent, Iris, 641, 651, 687
Berko, Jean, 662, 673
Bermúdez-Otero, Ricardo, 679
Bernhardt, B. H., 578, 591
Beskow, Jonas, 493

Bessell, Nicola, 249–50, 278
Bickel, Balthasar, 69
bilingualism, 572, 594
  and loanword adaptation, 755–7, 763–4, 773
Birch, Barbara, 760
Bird, Steven, 29, 37n, 205
Bishop, Jason, 307, 308
Black, Alan W., 518
blank-specification, 16
bleeding, 42, 43, 62–4, 742
  definition, 41, 65n
Blevins, Juliette, 135n, 164, 184, 325, 330–2, 336, 338, 340, 341, 345, 347n, 690n
blinking (boundary marking), 698, 699f, 700–1
Bloch, Bernard, 133n
blocking, 10–11, 12–16, 26, 51, 78, 80–1, 443–8, 476, 480n–1n, 482n
  definition, 7, 47–8
  harmony systems, 245–7, 248, 249, 254, 256, 258–9, 270–4, 280, 289n
  underapplication, 47–54
Bloomfield, Leonard, 135n, 351, 363, 364, 572
Blumenfeld, Lev, 267–8
Blumstein, Sheila, 310
Blust, Robert, 219
Boas, Franz, 453
Boberg, Charles, 433n
Bodomo, Adams, 721n
Boersma, Paul, 266, 277, 290n, 304, 391, 408, 413, 414, 419, 426, 591, 593, 621, 670, 689n
Bollt, Alex, 728, 748
Booij, Geert, 85, 100–1, 553, 554, 690n
Borno, 221
Bortfeld, H., 594
Boston University Radio News Corpus, 376
Botne, Robert, 739–40, 750n
boundaries
  in harmony systems, 282–4
  morphology-phonology interaction, 68–83, 224
  phrase boundary marking, 155–6, 224, 457–68, 484n, 508, 509, 522–3, 698, 699f, 700–1
  and syllable structure, 168–9, 174, 522–3, 600–2, 609, 623–4
brace notation, 9, 11, 12, 38n
bracket erasure, 82–3
Braconnier, Cassian, 212–13
brain, 36, 398, 704–5
Brandão de Carvalho, Joachim, 565
Brannen, K., 582
Brault, Frédérick, 761
Brazilian Portuguese, 739, 742
Breeze, Mary, 242
Bregman, Albert S., 704
Brennan, Mary, 692, 711
Brentari, Diane, 692, 693, 695, 696, 697, 700, 701, 702, 704, 714, 716–17, 720n
Bresnan, Joan, 602
British English, 495–8, 502, 775
Broe, Michael, 79, 262, 344
Bromberger, Sylvian, 38n, 41
Broselow, Ellen, 56, 113, 134n, 247
Browman, Catherine P., 166, 263, 325, 328, 369, 405, 432, 577
Brown, Gillian, 737
Brown, Heather, 764
Brown, Roger W., 648
Bruce, Gösta, 499, 510–13, 515
Bruening, Benjamin, 96, 731
Buchan Scots, 272
Buckeye Corpus of Conversational Speech, 375–6, 378, 381, 430
Buckley, Eugene, 747
Buckley, M., 583
Buder, Eugene H., 575, 592
Bueno-Holle, Juan, 293
Bukusu, 243
Bulgarian, 78
Bunuba, 300
Burger, Lisa, 281
Byarushengo, E., 233
Bybee, Joan, 88, 184, 230, 328–30, 333, 334–5, 347n, 378, 393, 405, 427, 428, 429
Byrd, Dani, 249, 518

Cahill, Mike, 217, 232
Cahuilla, 123, 137n–8n
Calabrese, Andrea, 3, 24, 255, 271
Calabrese Italian corpus, 764, 765t, 767–8
Calder, Jonathan, 613, 614
Campbell, Nick, 518
Campidanian Sardinian, 135n

Canadian English, 759, 767, 769
Canadian French, 582, 583, 679–80
Canalis, Stefano, 195n, 272
candidate chains, 31, 35
canonical downstep, definition, 222
Cantineau, Jean, 372n
Cao, Jianfen, 528
Capanahua, 245–6, 281
Cara (Tera), 15, 38n
Caramazza, Alfonso, 637
Carlson, Rolf, 599
Carson-Berndsen, Julie, 609
Carter, A., 585
Carter, Hazel, 500–1, 527
Carter, M., 773
Casali, Roderic F., 272, 281
case marking, 203–4
Cassimjee, Farida, 206, 440, 444, 449, 480n
Catalan, 294, 306, 585
causative suffixes, 77–8
Cayuvava, 144
CCC model (Evolutionary Phonology), 330–2
Cebrian, Julie, 737
Cedergren, Henrietta, 404, 527, 528
Chafe, Wallace, 33, 235
Chaha, 244
chain shifts, 46, 369–70, 573
Chambers, Kyle E., 738
Chamorro (Tjamoro), 94
Chan, Marjorie K.-M., 228–9, 523, 525, 526
Chang, Nien-chuang T., 357
Chao, Yuen Ren, 488, 515–16, 523, 528, 532n
Charette, Monik, 555, 557, 558–9, 562, 569n, 582
Charles-Luce, Jan, 306, 664, 669
Chase, Chris, 704
Chatino, 200
CHATR (computer software), 518
Chen, Matthew Y., 104, 133, 198, 460
Chen, Yiya, 504, 508, 525
Chengdu, 357
Cherry, E. C., 629n
Chicano English, 407, 409, 422f, 429
Chichewa, 217
Chickasaw, 151, 152, 156, 159
Chida, Shuntaro, 231
child-directed talk, 499, 584–5, 588, 701
CHILDES database, 579, 593
children, as research subjects, 287, 591–4 *see also* acquisition
ChiMwiini (Swahili dialect), 458–60
Chinese
  Mandarin Chinese
    Beijinghua (Chinese dialect), 501f, 504, 523, 528
    Mandarin *see* Putonghua (PRC Standard Chinese)
    Putonghua (PRC Standard Chinese)
      contrast reduction, 300
      loanword adaptation, 761
      syllable structure, 514
      tones, 208–9, 225–6, 357, 493–4, 501f, 502, 504, 516, 523, 528
  Min Chinese
    Fuzhou, 299
  Wu Chinese
    Shanghai, 225–6, 453, 525–6, 527–8
    Tangxi, 525, 527
    Wuxi, 227–9, 525–6
  Yue Chinese
    Cantonese
      morphology, 523, 527
      near mergers, 306, 308
      phonological contrasts, 291, 292–3, 521–4
Chingp'o (Jingpho), 87
Chomsky, Noam, 5–6, 7, 8, 12, 23, 31, 38n, 48, 49, 66n, 68, 82, 104, 134n, 206, 239, 304, 320, 349, 353–4, 356–7, 363, 366, 370–1, 428, 435, 440, 454, 534, 573, 597, 600, 602, 603, 604, 631, 636, 637, 660n
Chu, Shinying, 653
Chuave, 230–1
Chumash, 101, 242, 261
Church, B. A., 396
Churma, Don, 735
Chuvash, 122–3, 162
Clahsen, Harald, 673
Clark, Mary Morris, 498
clash (stress clash), 144, 150, 151
classifier handshapes (sign languages), 716–17
clauses, 452–4
C'Lela, 284
Clements, G. N., 19, 20–1, 24, 52, 177, 198, 209, 214–15, 222, 233, 239, 257, 269, 279, 297, 304, 544, 580, 612

clitics, 453–4, 586, 587, 592
  in sign languages, 698–700
  tones, 219
  and vowel harmony, 284–6
close-copy stylization, 495–7
CLPF corpus (Dutch), 582
clusters, 184, 405, 551–7, 568n, 581–2, 600–2, 615, 651
co-speech gestures, 701, 719
coarticulation, 325, 330–1, 332, 333–4, 343, 512–13
  and harmony systems, 266, 324
cochlear implant recipients, 650
Coda Condition, 111, 135n
coda consonants, 105–12, 128–9, 137n–9n, 190, 334
codas
  constraints, 188–92, 296–7
  definition, 134n, 167
  in harmony systems, 272–3
  in language acquisition, 580–1, 582–5, 588–90
  modeling of, 170–1, 181, 184–7, 337–8, 536–7, 546, 552–3, 555, 556–63, 566n
coerced weight, definition, 139n
Coetzee, Andries W., 406, 408, 410, 411, 414, 429, 430, 650, 667
cognitive theories *see* psychological realism
Cohen, Antonie, 495
Cohen, Michael M., 649
Cohen, P., 599
Cohn, Abigail, 145, 363
Cole, Jennifer, 273, 287, 290n, 605, 738
Cole, R. A., 396
Coleman, John, 29, 37n, 368, 504, 604, 610, 611, 615
Collier, René, 495, 505
commutation tests, 313
comparatives, 85
competence, 632–7, 646–51, 657–9, 660n
  gestural competence, 701
complement notation, 10–11
Complementary Learning System, 398
complete vowel harmony, 256
Complex NP Constraint, 7
complexity (head–dependent relationships), 535, 538–9, 547, 552–7, 566n, 568n–9n
complexity (sign languages), 718–19
compound constructions, 83–90, 99–102
compression algorithms, 399–400, 630n
compression (tones), definition, 221
computational phonology, 3, 4, 353, 592–3, 596–628
  agent-based simulations, 333
  algorithms, 399–400, 415–16, 419–20, 621, 630n, 634, 635, 665–72, 681–2, 685–8, 689n
  analysis-synthesis method, 495–7, 510, 518, 599–600
  artificial language-learning tasks, 343–4, 391
  Automatic Speech Recognition systems, 380, 383–4, 599–600, 610–11, 625, 627, 630n
  context-free languages, 614–16f
  finite state automata, 184–7, 603–11, 614–15, 618–21, 629n, 666, 686
  "generate and test" method, 611–12
  influence of information theory, 596–603
  learner simulations, 316, 396–7, 398, 399–400, 415–23, 431–3, 434n, 594–5, 621, 622–5, 663, 665–72, 685–8
  and lexical storage, 626–7
  n-gram models, 669
  software, 416, 417, 419–20, 423–5, 493, 518, 593, 599–600, 605
    development of, 628
  speech synthesis, 627
  spreading activation networks, 634–5
  text-to-speech systems, 600
  X-JToBI tagging, 506–7, 508, 515, 516
conditioning regularities, definition, 195n
connectionist models, 622–5, 630n
Connine, Cynthia M., 395
consonants
  articulation, 503
  clusters, 184, 405, 551–7, 568n, 581–2, 600–2, 615, 651
  consonant harmony, 241–5, 258–9, 260–3, 289n, 587
    typology, 260–1
    vowel-consonant harmony, 245–50, 263–8, 270–1, 272–3, 274–8, 281, 282–4
  contrast reduction, 292–3, 294, 295

deletion, 55, 56, 405, 406–7, 409–13, 421–3, 424–6, 429–30, 431–3, 433n–4n
input/output relations, 28, 32, 35
learner simulations, 416–20
onset restrictions, 20–1, 30–1, 514, 528–9
place assimilation, 326–7, 331–2, 404–5
conspiracy theories, 16–17, 18, 27, 33–5
in acquisition, 576
blocking rules, 52
feeding/-bleeding relationships, 57–8, 63, 66n, 67n
triggering conditions, 53–4
constituent structure hypotheses, 168–71, 178–80, 181–3, 194, 195n, 366, 435–78, 480n–4n, 534
in syntax, 183, 195n
Constraint Demotion Algorithm, 621
constraints, 6–7, 13–15, 25–37, 409–14, 611–21 *see also* templates
acquisition, 574, 576, 578–80, 583–7, 590–1, 621, 665–72, 675–8, 685–8, 689n–90n, 737–8
alignment constraints, 264–6, 450
complexity, 538–9
contrast reduction, 293–4, 298–304, 309
definition, 4–5, 30
domain-specific constraints, 285, 435–7, 441–78
edge constraints, 110–12, 117–18, 135n, 159, 184–7, 189, 442–7, 449–51, 456, 458–62, 466–7, 469–70, 471, 476–8, 481n, 483n, 552–7
evaluative constraints, 16–18
faithfulness constraints, 82, 95, 101–2, 182, 267, 269, 271, 317n, 354, 410–12, 451, 470, 481n, 668, 680, 685
gradience, 639–44, 669–72
harmony systems, 264–8
head–dependent relationships, 538–40
if-then constraints, 16–18
indexation, 71, 72–3, 75
language games, 726–36, 737–41, 744–5, 748
lengthening, 49, 106–7, 116–17, 441–3, 480n
licensing conditions, 284, 296–9, 317n, 535, 539, 540, 550, 557–63, 569n–70n, 587, 652
loanword adaptation, 751–78
markedness, 15, 342, 410–11, 451, 455, 468–78, 483n, 484n, 552, 583, 668, 745–7
Match Theory, 439–40
metrical stress theory, 146–51, 155–6
Morpheme Structure Constraints (MSPs), 734–6, 738, 748
No-Crossing Constraint (NCC), 258, 259, 277
paradigm contrast, 78–80
paradigm uniformity, 76–8
Preservation Principle, 763, 774–6
ranking of, 181, 270–1, 277, 298–9, 301–2, 367, 408–23, 425–7, 434n, 448–50, 466–7, 469, 481n–2n, 539–40, 565, 583, 591, 594–5, 617–18, 626, 665–72, 685, 689n–90n, 723–4, 726–7, 734
underspecification, 20, 541, 587, 611–12
universals, 320–2, 336–45, 621
Weight-by-Position rule, 105–8, 135n, 138n, 139n–40n
Weight-to-Stress constraint, 128
well-formedness *see* well-formedness
word formation, 83–90
Constraints in Phonology Project (CoPho), 751–5, 757–78
methodology, 757–61, 776–7
consultants (research participants), 650, 759, 776–7
contact situations, 754–5
context dependent weight, 127–9, 135n–6n, 138n, 139n–40n
context-free models, 614–16f
continuancy, definition, 564
continuous harmony, definition, 240
contour tones, 201, 208–10, 215, 300
contrast, definition, 291, 304–5, 310
contrast enhancement, 333–4
contrast preservation, 84–5, 299–304, 309–16
in harmony systems, 271
contrast reduction, 291–316, 375–6
articulation, 300, 312, 328
theories, 295–304, 308–16, 321–3, 335, 364
typology of, 292–3
Contreras, Helen, 170
cooccurrence, 171
Coordinate Structure Constraint, 7

CoPho (Constraints in Phonology) project, 751–5, 757–78
  methodology, 757–61, 776–7
cophonologies, 71–2, 73, 81, 82, 99
  definition, 71
Coppola, Marie, 711
coronal consonants, 11, 12, 69
  harmony systems, 241–3, 260–1
  markedness rules, 13
Corpus of Spontaneous Japanese, 515, 528
corpus studies, 374–400, 656, 751–78
  methodology, 380–8, 579, 582, 588, 592–3, 627–8, 752–5, 757–61, 767, 776–7
Correspondence Theory, 31, 262–3, 278, 289n, 290n, 450–1, 452–4, 479n, 482n
  Base-Reduplicant Correspondence Theory, 95, 101–2
Coulter, Geoffrey, 702
counterbleeding, 42, 58–9, 62–4, 65n, 66n–7n, 679, 680, 742, 744
  definition, 41, 58, 62
counterfeeding, 27, 42, 43–7, 65n, 679
  definition, 41, 43, 44–5
Coussé, E., 383
covert contrast, 309–16, 364, 575–6, 583–5, 592
Craik, F. I. M., 396
Crasborn, Onno, 696–7
Crawford, Penny, 306
creaky voice, 493, 494, 522, 525
Cree, 378–9
Creek (Muskogee), 94
creolization, sign languages, 711, 712
Croatian, 515
cross-derivational feeding, 60–2
cross-dialectal comparisons, 409–13, 421–3
cross-level rules, definition, 28
cross-linguistic comparisons
  acquisition, 576–8, 583–5
  contrast reduction, 293–5, 300, 306–7, 313–16
  frequency effects, 393
  intonation, 498, 513–14
  language games, 735–6
  loanword adaptation, 761–78
  morphology-phonology interaction, 83–90, 702–3
  nasal harmony, 246, 271
  sign languages, 700–1, 702–3, 705–8, 721n
  sound change, 336–45, 373n
  syntax-phonology interface, 448–50, 456–7, 463–4, 467–8, 472
  tones, 527–30
  VOT (voice onset time), 359–60
  vowel harmony, 254, 255–6
  well-formedness judgments, 641, 648–9, 651
Crossley, Laurinda, 700
Crossover Constraint, 13–14
Crosswhite, Katherine, 78, 294, 327
crowding, 156
Crowhurst, Megan, 114–16, 136n
Cruttenden, A., 576
Cuban Spanish, 334
culminativity, 232–3
Cummins, Robert, 632, 635
Cuna, 735–6, 742
Curtin, Suzanne, 689n
Curtis, Emily, 123, 131, 137n
Cutler, Anne, 284, 634
CV phonology, 489, 544, 563–4, 567n
cyclic rules, 605
Cyran, Eugeniusz, 556

/d/ deletion, 405, 406–7, 409–13, 421–3, 424–6, 429–30, 431–3, 433n–4n
Dadibi, 200, 231
Daelemans, Walter, 392, 400
Dagaare, 721n
Dakota (Sioux), 69
Daly, John P., 210–11
Danesi, Marcel, 763, 764
Danis, Catalina, 740
Danish, 135n
dash-factoring notation, 11–12
data analysis, 384–8
data collection, 201–2, 286, 287, 380–2, 520–1, 582, 593–4, 652–5, 759–60
  story-telling, 378–9
Dauer, Rebecca M., 527
Davidson, Lisa, 639, 649, 654–5, 773
Davies, H. J., 153
Davis, Barbara L., 499, 577
Davis, Stuart, 117, 118, 120, 122–3, 124, 135n, 137n, 139n, 170, 171, 248, 270, 281, 730, 739–40, 750n

de Boer, Bart, 336
de Jong, Ken, 133, 312
de Lacy, Paul, 337, 339, 341–2, 461
de Pijper, Jan Roelof, 495–7
de Rijk, Rudolf, 46
de Vaan, L., 386
debuccalization, 298
decision-making processes, 642–4
Declarative Phonology, 29–30, 37n, 612–16, 621
declination line (intonation), 497
default specification, 20
degemination, 57–8, 63, 66n
degenerate feet (stress), 123, 158–9
degenerate syllables, 552–5, 568n
Dehaene-Lambertz, Ghislaine, 649
Dehé, Nicole, 453
delabialization, 46–7
Delattre, Pierre, 588
deletion, 38n, 39n, 55, 56, 405, 406–7, 409–13, 429–30, 431–3, 433n–4n
  apocope, 27, 28
  counterfeeding, 45, 49–50, 57–8
  tones, 224–5
Dell, François, 176–7, 431, 504–5, 552, 557, 568n
Dell, Gary, 281, 652, 653, 738
Demuth, Katherine, 578, 582, 583–4, 585, 586, 588, 589, 591, 592
dental consonants, 242
Denwood, Ann, 555
dependency-based relationships *see* head–dependent relationships
Déreau, Léon, 262
derivational relations, 16–18, 25, 31–2, 35–6
derived sonority models, 180–3
descriptive grammars, 205
determiners, 586
Dettweiler, Stephen, 284
devoicing
  environment-related effects, 17–18, 29–30, 58–9
  learner simulations, 416–20
  universal principles, 24, 337–9
  vowels, 38n, 39n, 527–8
diachronic sound patterns *see* sound change
Diehl, Randy L., 309–12, 326–7, 737, 744, 745–6
Diffloth, Gerard, 93
Dilley, L. C., 375–6
diminutives, 76–7, 97
Dinnsen, Daniel A., 24, 32, 306, 307, 369
Diola Fogny (Jola-Fonyi), 49–50
Dioula, 212–13
direct mapping, 27–8, 39n
discourse effects, 379, 452–3, 505, 526
discrimination tasks, 648–9, 650
disjunctive blocking, 48–9
disordered speech, 591, 653–4
Dispersion Principle, 177
Dispersion Theory, 302–4, 305, 745
dissimilation, 69, 78–80
  and sound change, 325, 331
  tones, 219–20
distance harmony, definition, 240
distinctive weight, definition, 139n
Dixon, Robert M. W., 702
do-something-except-when blocking, 52–4
Dogon, 231
Dohlus, Katin, 752, 778n
Dom, 231
domain, definition, 480n
domain-specific constraints, 285, 435–7, 441–78, 480n
dominant hand (sign language), 721n
Donohue, Mark, 230–1, 232–3
Donovan, R. E., 610
dorsal consonants, 243–4, 261
dos Santos, C., 583, 587
doubling, 77–8 *see also* reduplication
Doubly Filled Comp Filter, 603
Downing, Bruce, 77, 91, 92, 93, 96, 117–18, 198, 215
downstep, 222–3, 224, 460–1, 473–8, 509, 517–21
  definition, 376, 517
Dras, Mark, 336
Dresher, B. Elan, 291, 537, 538–9, 675, 677, 678
Duke of York derivations, 47, 48–9
Duplication Problem, 52
Duponceau, Peter, 134n
Dupoux, Emmanuel, 647, 648–9
duration *see* quantity
Dutch
  acquisition, 579, 581, 582, 588, 664, 672, 673, 674, 675–6, 681, 682, 690n

Dutch (*Cont'd*)
contrast reduction, 292, 308, 335, 393
corpus studies, 375, 379, 381, 382, 390
information structure, 379
intonation, 495, 497–8, 514
lexicon, 392
morphology, 85, 100–1, 390, 392
reduction, 378, 394
sociolinguistic variation, 379
stress, 160, 579
syllable structure, 188, 393, 552–4, 555
t/d deletion, 430

early words *see* acquisition
Eastern Andalusian Spanish, 307
Eberhard, David, 217
Ebert, Karen, 244, 360
edge constraints, 110–12, 117–18, 159, 184–7, 189, 442–7, 449–51, 456, 458–62, 466–7, 469–70, 471, 476–8, 481n, 483n, 552–7
Edmondson, Jerold A., 200
Edmondson, T., 293
Educated Singapore English, 57–8, 63, 66n, 67n
Edwards, J., 588
Efere, Emmanuel Efereala, 226
Eisner, Jason, 666
Elenbaas, Nine, 149
Elfner, Emily, 470, 741
elicitation techniques, 201–2, 286, 287, 380–2, 520–1, 582, 593–4, 652–5, 759–60
story-telling, 378–9
Ellis, Shirley, 731
Ellison, T. Mark, 609, 610f, 618–20, 666
Elman, Jeffrey, 622, 634
Elmedlaoui, Mohamed, 176–7, 504–5, 557
Elordieta, Gorka, 473–8, 479n, 484n
Elsewhere Condition, 48–9, 51
Embick, David, 479n
Emeneau, M. B., 294
Emmorey, Karen, 709
Emonds, Joseph, 452
emphasis harmony, 247–8, 270, 271–2, 281, 289n, 290n
empirical adequacy, definition, 37n
empty-headed constituents, 545, 547, 548–57, 558–63, 567n, 568n–9n, 582
encliticization, 453–4
endangered languages, 378–9
Endress, Ansgar, 728
Engberg-Pederson, Elisabeth, 716
Engenni, 221
England, Nora, 70–1
English *see also* specific varieties of English, e.g. Singapore English
acquisition, 579–80, 582, 583–5, 586, 587, 588, 589–90, 592, 664, 673–4, 683, 686–7
bilingual acquisition, 594
consonants
articulatory overlap, 325
clusters, 405, 552, 615
deletion, 55, 56, 405, 406–7, 409–13, 421–3, 424–6, 429–30, 431–3, 433n–4n
devoicing, 60–2
flaps, 191–2, 306, 312–13, 394, 395
onset restrictions, 20–1, 344
place assimilation, 404–5, 407
rhotic consonants, 376–7, 384–6, 756, 774–5
stops, 755–6
/t~d/ contrast, 293, 306
Velar Softening rule, 55
contrast reduction, 291–2, 306, 310–11, 312–13, 314, 375–6
corpora, 375–7, 378, 380–2, 387–8
glides, 754
intonation, 376, 470, 495–8, 502, 524, 527
language games, 726–7, 729–30, 731, 737–8, 740, 741, 744, 748
lexically conditioned variation, 428–30, 431–3
loanword adaptation, 751, 754, 759, 760, 761, 767, 769, 773, 777
morphology, 60–2, 85–6, 89, 99
reduction, 378, 453
reduplication, 95
semivowels, 325
stress, 55, 160, 390, 428–9
syllable structure, 165, 184–7, 191–2, 310–11, 514, 555–6
Trisyllabic Laxing rule, 15, 27, 48–9
vowels, 60–2, 134n, 160
formant values, 361, 362f
nasalization, 363

environment-related effects, 164–5, 167–8, 190–2, 386, 492
  blocking, 51, 78, 80–1, 272–3
  contrast reduction, 293–4
  feeding/bleeding relationships, 45, 58–9, 63
  misperception of, 323–6, 643–4
Epena Pedee, 245, 246, 289n
epenthesis, 62–3, 307, 608–9, 746
  blocking conditions, 14, 57–8, 476, 483n
  ordering, 16, 87–8
episodic models, 395–7, 627
Epps, Patience, 231
Ernestus, Mirjam, 308, 375, 378, 379, 381, 383, 386–7, 390, 394, 395, 427, 681
errors
  perceptual errors, 322–6, 643–4, 647–51, 660n
  short-term memory, 655
  speaker errors, 281, 287, 651–5, 660n
Esimbi, 217
Esper, Erwin, 737–8, 750n
Estonian, 121
Etung, 293
evaluation, 1–5, 7–9, 16–18, 31–7, 37n, 66n
Evans, Nicolas, 249
Event Visibility Hypothesis, 714–16f
Evolutionary Phonology, 330–2, 336, 341–2, 345
Ewen, Colin J., 543, 548
exemplar-based models, 333–6, 361, 368, 388–400, 627
existential quantifier, 39n
expansion (tones), definition, 221
extrametricality, 128–9, 140n, 235–6
extrasyllabicity, 552–3, 554–5, 569n

f_0 *see* fundamental frequency
Fabiano-Smith, Leah, 56
faithfulness constraints, 82, 95, 101–2, 182, 267, 269, 271, 317n, 354, 410–12, 451, 470, 481n, 668, 680, 685
falsification, 41, 64–5, 541–2, 552
Fant, Gunnar, 355–6, 366
Fante (Akan dialect), 52
Faroese, 744
Farwell, Carol B., 576, 577
Fasu, 231
faucal harmony (retraction harmony), 249–50, 278
feature-coefficient variable, 9–10
feature-filling patterns, 257, 612
feature geometry, 8, 268–9, 297–8, 320, 544, 581, 612–16
features
  alignment constraints, 264–5
  definition, 356, 372n
  feature economy, 304
  in harmony systems, 258–9, 265
  limitations of, 169, 238–9, 355–8
  matrices, 8, 597, 598f, 613–16
  sign languages, 691–2, 693–5, 714–16f
fed counterfeeding, 45, 46–7, 48–9
Fee, E. Jane, 579
feeding, 27, 41, 42, 43, 60–2, 65n *see also* counterfeeding
feet (stress feet), 469–70, 489–91, 574–5, 585
  definition, 145–6, 437
  and harmony systems, 267–8, 283–4
  metrical stress theory, 145–54
Feng, Gary, 733–4
Ferguson, Charles A., 576, 577
Féry, Caroline, 172, 196n, 482n
Fidelholtz, James, 429
field work, 201–2, 311, 313, 572, 574, 776–7
Fijian, 159–60, 753
Fikkert, Paula, 579, 673
filtering, definition, 7
final lowering (tones), 517, 518–19
fingerspelling (sign language), 696–7, 700
finite state automata, 184–7, 603–11, 614–15, 618–21, 629n, 666, 686
Finnish
  assibilation, 14–15, 51, 80
  consonants, 28, 32, 35
  lexicon, 626–7
  morphology-phonology interaction, 287, 407
  prosodic features, 131
  vowel harmony, 287, 290n
Finnish Sign Language (FinSL), 696
Firchow, Irwin B., 232
Firchow, Jacqueline, 232
first-language acquisition *see* acquisition
Firth, J. R., 488
Fischer-Jørgensen, Eli, 164, 168, 189
Fisher, Cynthia, 738

Fisher, R. A., 671
Fisher, Ryan, 767
Fisher, W., 380
Fitzgerald, Colleen, 129, 140n
fixed-level phonology, 27–8
fixed stress systems, 142, 143, 144–5
flaps
  contrast reduction, 191–2, 293, 306, 312–13
  in harmony systems, 271
flat structure models, 174
Flemming, Edward, 271, 283, 302, 311, 361, 362f, 434n
floating tones, 210–12, 238, 502, 511–12
focus, 198, 528–30
  counterbleeding, 59
  counterfeeding, 46–7, 48–9
Fodor, Jerry A., 622, 637
foot (stress foot) *see* feet (stress feet)
Force Phrase, definition, 452
Ford, Kevin C., 19, 227
formalism
  importance of, 1–2
  limitations of, 15–16, 29–30, 193–4
  properties of statements, 5, 597–9
  see also modeling; notation; well-formedness
Fougeron, Cécile, 308, 431, 528
Fourakis, Marios, 307
Fowler, Carol A., 524
Fox, Robert A., 306, 312
Frank, R., 620
Franks, Steven, 143
Freitas, Maria João, 584
French, 190–1, 435–6, 479n, 581, 582, 775
  acquisition, 581, 582, 583, 586, 679–80
    bilingual acquisition, 594
  loanword adaptation, 751, 755–6, 759, 761, 765, 768–9, 770–1, 772
  schwa-deletion, 308, 394–5, 430–1, 558–9, 569n
  syllable structure, 547, 649
  syntax-phonology interface, 435–6, 479n
frequency
  and language acquisition, 580–1, 583, 584–5, 588–90, 669
  and pronunciation, 377, 394–7, 429–30, 653–4
  and sound change, 329, 332, 336
  and word recognition, 393, 395
Friederici, Angela D., 664, 737
Friesner, Michael, 479n
Frisch, Stefan, 79, 262, 344, 391, 655
Frishberg, Nancy, 709
Friulian, 314–15
Fromkin, Victoria, 210
Frota, Sónia, 584
Fry, D. B., 493
Fudge, Eric, 164, 171, 555, 580
Fujimura, Osamu, 312
Fujisaki, Hiroya, 487, 510, 517–19, 520
Fula, 26, 121–2, 766
  vowel harmony, 254–5, 280
Fulmer, S. Lee, 87
fundamental frequency, 490f, 491, 493–5, 508–9, 510, 520
  and downstep, 475
  harmonics, 131
Fung, Roxana Suk-Yee, 523
Furby, Christine, 148–9t
Fuuta Tooro Pulaar, 88 *see also* Fula

Gaelic, 246, 270–1
Gafos, Adamantinos, 92, 249, 261, 271, 275, 288, 335
Gahl, Susanne, 377
games *see* language games
Ganda (Luganda) *see* Luganda (Ganda)
gapping, 85, 259–60, 289n
Garawa, 148, 150
Gårding, Eva, 493
Garrett, Andrew, 325, 336, 341
Garrod, S., 386
Gaskell, M. Gareth, 633
Gathercole, Susan, 655–6
Gazdar, Gerald, 602, 603, 612
geminate consonants, 14
  and contrast reduction, 297
  formal description of, 175, 190
  learner simulations, 416–20
  in moraic phonology, 105–6, 119–27, 129, 134n, 137n–40n
gemination (sign languages), 721n
gender, grammatical *see* grammatical gender
gender marking, 199–200, 204
  diminutives, 76–7
Generalized Phrase Structure Grammar (GPSG), 612, 614

generative grammar, 4, 5–7, 596–603, 604–5, 639
generative phonology
  influence of information theory, 596–603
  phonetics-phonology interface, 320, 340–1, 354, 370–1
  representation of quantity, 104–8
  scope of inquiry, 2, 341–2
  syllable structure, 172–5
  transformational grammar, 604–5
  and variability, 388–9
genitives, 204, 210
Geoffrion, Louis-Philippe, 759
Georgian, 557
Gerfen, Chip, 307
Gerken, LouAnn, 585, 586, 594, 728, 748
German, 462–3, 775
  acquisition, 582, 673
  contrast reduction, 306
  devoicing, 17–18, 29–30, 59, 63, 67n, 335
  loanword adaptation, 772
  morphology, 78–9
  prosodic features, 454, 462–3, 482n–3n, 519–20
  spirantization, 59
  syntax-phonology interface, 454, 462–3, 482n–3n
German Sign Language, 721n
gerundives, 96
gestures (articulatory gestures) *see* articulation
Ghazeli, Salem, 247
Ghini, Mirco, 439, 468
Gibson, Lorna F., 231
Gick, Bryan, 288
Gilbert, J., 578
Gildea, Daniel, 686
Gimira (Benchnon), 242–3
Gimson, A. C., 375
Girón Higuita, Jesús María, 232
Giryama, 223–4
glides, 45, 272, 290n, 754
  definition, 169
globality, 6–7, 33–5
  definition, 5
glossematics, 195n
glottal consonants, 138n
Glover, Bonnie, 15
Goad, Heather, 253, 579, 581, 582, 583, 587
Godfrey, J., 377, 380–1
Goedemans, Rob, 118, 136n
Goeman, A. C. M., 430, 434n
Golder, Scott, 726
Goldilocks Rule, 728, 748
Goldin-Meadow, Susan, 711
Goldinger, Stephen D., 368, 396–7, 398, 627
Goldrick, Matthew, 308, 362–3, 652, 653, 654, 655
Goldsmith, John A., 18–19, 28, 36, 155, 164, 180, 190, 206, 208, 219–20, 233, 237, 257, 258, 366–7, 483n, 488, 507, 574, 602, 683, 691
Goldstein, Louis M., 166, 263, 323, 325, 328, 369, 405, 432, 503, 524, 577
Goldvarb (computer software), 423–5
Gollan, Tamar H., 709
Gombe dialect (Pulaar), 88
Gomez-Imbert, Elsa, 229
Gordon, E., 376, 381–2
Gordon, Matthew, 117, 130–2, 134, 156, 157, 159, 162, 290n, 300
Gosiute Shoshone, 159
Goulet, Marie-Josée, 767
Gouskova, Maria, 347n
Gout, Ariel, 649
government phonology, 533–66, 566n–70n
  constraints, 538–40, 557–63
  definitions, 533–9
  empty-headed constituents, 545, 547, 548–57, 558–63, 567n, 568n–9n, 582
  intrasegmental relations, 537–8, 541–5
  notation, 535–7
gradience, 639–44, 649–51, 652, 653–5, 669–72, 684–5
Gradual Learning Algorithm (GLA), 415–16, 419–20, 621, 670–1, 685, 689n
grammar competition (interference), 56–7, 488
grammar dependence, definition, 73
grammars, 34–5, 205 *see also* specific theories of grammar, e.g. Universal Grammar (UG)
grammatical gender *see* gender marking
Grammont, Maurice, 323
grandfather effects, 52–3
Granström, Björn, 599
Grassfields Bantu languages, 211

Greek, 136n, 325
  acquisition, 673
  prosodic features, 100, 514, 528
Greenberg, Joseph H., 214, 641, 660n
Greenlandic, West, 15–16
Gregerson, Kenneth J., 200
Grice, Martine, 513–14
grid-based models, 156–62
Grønnum, Nina, 519, 520
Grose, Donovan, 717
Grosjean, François, 755
Grossberg, S., 398, 399
Grounded Phonology, 24–5, 33, 270, 280, 341, 357–8, 456
Grundzüge der Phonologie (1939), 348–9, 352, 355, 371n–2n
Guaraní, 275–6, 283–4, 776–7
Guere, 27
Guerra, Anne-Marie Currie, 710
Guerssel, Mohamed, 14, 29
Guimarães, Maximiliano, 739, 742, 744
Guion, Susan G., 326, 327, 347n, 687
Gujarati, 339, 347n–8n
Gupta, Abha, 123
Gussenhoven, Carlos, 155, 156, 198, 235, 477, 498, 513, 520, 521, 529
Gussman, Edmund, 178, 557
Guy, Gregory R., 406–7, 409–10, 433n

"H & H theory", 327, 332, 434n, 516
Hahn, Ulrike, 641, 669
Haitian Creole, 84
Hakha Lai, 202, 231
Hale, Kenneth, 45, 62, 91, 110
Hale, Mark, 3, 36, 368, 466
Halh, 252–3, 259, 273
Hall, Beatrice L., 220
Hall, R. M. R., 220
Halle, Morris, 8, 12, 16, 19–20, 33, 38n, 41, 48, 49, 66n, 68, 82, 104, 134n, 206, 238–9, 304, 320, 355–7, 363, 366, 370–1, 428, 435, 573, 597, 599, 600, 604, 610, 626, 627, 629n, 631, 636, 660n, 662
Halliday, M. A. K., 527
Ham, William, 128–9
Hamann, Silke, 300
"hammock" stress systems, 144, 148, 158
Hammond, Michael, 730, 731
handshape (sign languages), 691, 692, 693–701, 706, 707, 715–19, 720n, 721n
Hankamer, Jorge, 81, 85, 175
Hansen, Benjamin B., 294
Hansen, Kenneth, 142
Hansen, L. E., 142
Hansson, Gunnar Ó., 241, 242, 243, 244, 249, 261, 262, 279, 281, 289n, 680
haplology, 85–6
Hara, Noriyo, 518
Haraguchi, Shosuke, 19, 488, 507
Harari, 250, 278
Hargus, Sharon, 74–6
Harmonic Grammar, 416–17, 426–7, 671
Harmonic Phonology, 28, 35
  globality, 33
harmonics, and tonal perception, 131
harmony systems, 240–88
  articulatory effects, 254–5, 261, 263, 266, 273, 290n
  autosegmental representation of, 256–61, 264–5, 277–8
  blocking, 245–7, 248, 249, 254, 256, 258–9, 270–4, 280, 289n
  consonant harmony, 241–5, 258–9, 260–3, 289n
  constraints, 264–8
  definition, 240
  directionality, 279–82, 289n
  markedness, 277
  morphological domain, 279, 284–7, 289n, 290n
  parasitic harmony systems, 273
  phonological domain, 282–4
  root control, 220, 242–3, 279–81, 284–6
  spreading, 257, 258–60, 264–6, 270–1, 272, 279–80, 281
  and stress, 266, 267–8, 276, 283–4
  transparency, 259, 260, 274–7
  typology, 260–1
  vowel-consonant harmony, 245–50, 263–8, 270–1, 272–3, 274–8, 281, 282–4, 289n
  vowels, 52, 215–17, 218–19, 222–4, 237, 251–6, 257–8, 259, 266, 279–81, 335, 343, 669, 678–80, 690n
Harms, Phillip, 245, 289n

Harms, Robert T., 336
Harrington, Jonathan, 332
Harris, Alice, 98
Harris, James W., 171, 175, 682–3
Harris, John, 178, 376, 559
Harrison, K. David, 251, 281, 286, 336, 736
Harrison, Sheldon, 93
Harvey, Mark, 272
"hat pattern", 496, 498
Haugen, Einar, 312
Hausa, 74–5, 97–8, 124–5, 138n–9n, 293
Hawkins, Sarah, 398–9, 574–5
Hay, Jennifer B., 376–7, 384, 386
Haya, 233–4
Hayes, Bruce, 21, 24, 25, 62, 103, 105–8, 117, 119, 123, 133, 134n, 135n, 136n, 137n, 140n, 145–6, 151–2, 154, 159–60, 162–3, 287, 300, 301, 391, 408, 413, 414, 426, 434n, 468, 470, 621, 662, 667, 671, 672, 681, 683, 685, 686–7
Hays, D. G., 534
HCRC Map Task Corpus, 380
head–dependent relationships, 452–4, 472, 533–66, 566n–70n
  constraints, 538–40, 557–63
  definitions, 533–9
  empty-headed constituents, 545, 547, 548–57, 558–63, 567n, 568n–9n, 582
  inheritance hierarchies, 682–3
  intrasegmental relations, 537–8, 541–5
  intrasyllabic relations, 545–57
  notation, 535–7
  prosodic marking of, 235, 491–2
heads, 195n, 452, 567n–8n
  definition, 169–71, 533–9
Heath, Jeffrey, 231
heavy syllables, 107, 697
height harmony (vowels), 253–4, 269, 270, 276
Heinämäki, Orvakki, 287
Heinz, Jeffrey, 669
Henderson, Eugenir, 301–2, 317n
Hendricks, Sean, 92, 93
Herman, Rebecca, 520
Hermans, Ben, 105
Hertz, Susan R., 510
hiatus resolution, 34, 558–9, 561
Hidden Markov models (HMMs), 610–11, 625, 630n
Hildum, Donald C., 648, 737
Hindi, 113, 123, 153, 731
Hinskens, Frans, 430
Hintz, Diane, 145
Hirschberg, Julia, 524
Hjelmslev, Louis, 195n
Hoberman, Robert, 247, 290n
Hochberg, Judith, 574
Hock, Hans, 135n, 679
Hockett, Charles F., 134n, 524, 557, 604
Höhle, Barbara, 578
Hollich, George, 701
Holt, Lori L., 633
Hombert, Jean-Marie, 313, 735
Homer, Molly, 271
Homeric infixation (language game), 741, 749n
homesigning, 711–12
homophony, 78–80, 377
homorganic syncope, 22, 39n
Honda, Kiyoshi, 520
Hong Kong Sign Language (HKSL), 700
Hooper, Joan B., 37n, 174, 600–2, 677
Hopcroft, J. E., 605
horizontal assimilation, definition, 214
Houlihan, Kathleen, 15
Houston, Derek M., 705
Howard, Irwin, 13–14
Hua, 90
Hualde, José, 46, 276, 283
Huave, 87, 776
Hudson, Grover, 37n
Hulst, Harry van der, 127–8, 140n, 164, 188, 189, 198, 233, 258, 260, 537, 538–9, 544, 547, 557, 559, 562, 563, 565, 567n, 569n–70n, 693
Hulstaert, Gustav, 45
Hume, Elisabeth V., 126, 209, 237, 297, 299, 302, 679
Hungarian, 100, 140n
  vowel harmony, 286–7, 288, 335
  vowel/zero alternations, 549–51
Hunnicutt, M. S., 600
Hup, 231
Hura, Susan, 326–7, 331
Hyde, Brett, 118

Hyman, Larry M., 92, 126, 129–30, 136n, 195n, 199, 202, 204, 209–11, 214, 215, 216, 217, 218, 219, 222, 223, 224–5, 231, 233, 236, 238, 253, 279, 281, 285, 313, 323, 336, 489, 524, 530–1, 772
hyper- and hypoarticulation model, 327, 332, 434n, 516
hyperarticulation, 327–8, 347n
hypercorrective change, definition, 325
hypervariation, 726–7
hypocorrective change, 324–5, 343

iambic feet, 146, 470–1, 586
iambic stress systems, 151–2, 159–60
Icelandic, 78–9, 340, 744
iconicity
  sign languages, 692, 709–19, 721n
  tones, 204
identification tasks, 648
Identity Condition, 39n
identity effects, 273
Identity Principle, 95
Idsardi, William, 610, 649, 739
if-then constraints, 16–18, 25
IFA Dialog Video Corpus, 382
Igbo, 203, 204, 498
illocutionary clauses, 452–3
Ilocano (Ilokano), 407–8
Imdlawn Tashlhiyt Berber, 176–7, 557
indexical features, storage of, 396, 397, 398–9
Indo-European, 164–5, 167–8
Indonesian, 100, 145
Ineseño Chumash, 242, 258–9, 260, 281
infant studies, 287, 591–4, 664, 669, 695–6, 728
infixes, 85, 88–90
  language games, 726–7, 729–30, 732–3, 739–40, 741, 744–5, 746, 749n, 750n
  and reduplication, 94, 729–30
information gain trees (compression algorithms), 399–400
information structure, 379, 479n
information theory, influence on generative phonology, 596–603
Ingram, D., 578
inheritance hierarchies, 682–3
initial dactyl effect, definition, 471
Initial Tone Association Rule, 19
Inkelas, Sharon, 72, 73, 74, 75, 81, 82, 85, 92, 94, 95, 96, 97, 100, 101, 102, 216, 233, 439, 468–9, 675, 677
innateness *see* nativist theory (acquisition)
Innermost Wins principle, 74–6
input/output relations, 27–8, 32, 35, 451
insertion rules, 608–9
Institute for Perception Research (IPO) framework, 495, 497, 499–500
instrumental studies, 353, 358–61, 362f, 493–5, 520, 650, 660n
  harmony systems, 287
inter-speaker variation, 428–9, 516
interference, 56–7, 488
intergenerational language transmission, 322–3, 711–12 *see also* sound change
International Phonetic Alphabet (IPA), 350–2, 355, 357, 359–61, 362, 368, 372n, 498–9
International Phonetic Association (IPA), 349, 352, 372n
Interonset Licensing, 557, 569n
intonation, 437, 485–532
  anchoring, 500–1, 506–15
  and constituent structure, 441–3, 452–3, 454–5, 456
  definition, 485–6, 530–2
  downstep, 376, 517–21
  final lowering, 517, 518–19
  and fundamental frequency, 490f, 491, 493–5, 508–9
  metrical theories, 489–91, 527–30
  modeling of, 487–530
  perception cues, 503
  prominence projection, 527–30
  sandhi processes, 524–7
  in sign languages, 697–701, 721n
  and stress, 154–6, 470–2, 495–7, 527–30
  tonal space, 515–21
  transcription, 498–502
Intonational Phrase (ι), 441–3, 452–3, 454–5, 456, 470, 516–17, 518–19, 527
  definition, 437
  in sign languages, 697–701, 721n
intra-speaker variation, 428–9
intralevel rules, definition, 28
intrasegmental relations, 537–8, 541–5
intrasyllabic relations, 545–57
intrusive /r/, 376–7

inviolability, definition, 5
IPA *see* International Phonetic Alphabet (IPA); International Phonetic Association (IPA)
Isomorphism Hypothesis, 773–4
Israeli Sign Language, 721n
Isthmus Zapotec, 293
Italian
  loanword adaptation, 764, 765t, 767–8
  syntax-phonology interface, 457–8
  vowels, 217, 272
Italian Sign Language (LIS), 696, 721n
iteration, 12, 14
Itô, Junko, 72, 73, 81, 85, 100, 111, 190, 261, 297, 418, 420, 440, 446, 455, 456, 467, 468, 479n, 732
Iverson, Greg, 15
Izon, 226

Jackendoff, Ray, 203, 453
Jacobsen, William, 293
Jäger, Gerhard, 419
Jahr, Ernst Håkon, 726, 740
Jakobson, Roman, 355–6, 366, 573, 598f, 629n, 636
Jamaican English, 422f
Janas, Petr, 144
Jantunen, Tommi, 696, 697
Japanese
  acquisition, 584, 585
  geminate consonants, 418–20
  language games, 732
  loanword adaptation, 418–20, 468, 751, 756, 758, 765, 767, 772, 774–5
  modeling of, 608
  pitch accent, 232, 484n, 487, 488, 489, 490–1, 492, 502, 505–9, 510, 513
    and constituent structure, 435, 456, 460–1, 469, 472
    downstep, 474
    and morphology, 69–70, 71–3, 81, 100
    tonal space, 515–21
  quantity, 103, 104–5
  syllable structure, 111, 170, 489–91, 648–9
  syntax-phonology interface, 435, 456, 468, 469, 472
  vowels, 528–9
  word formation, 100
Japanese-English speakers, 488
Japanese Sign Language (JSL), 700
Jarosz, Gaja, 671, 676
Javanese, 94, 95
Javkin, H., 214
Jenison, Priscilla B., 232
Jenison, Scott D., 232
Jenkins, James J., 641, 660n
Jenner, Annette R., 704
Jensen, John, 14
Jessen, Michael, 340
Jingpho (Kachin), 87
Jita (Kijita), 77–8
Johnson, C. D., 604–5
Johnson, Keith, 302, 378, 395, 396, 397
Johnson, Mark, 583
Johnson, Stephanie, 361, 362f
Johore Malay, 271
Jola-Fonyi (Diola Fogny), 49–50
Jones, C., 602
Jongman, Allard, 247, 248, 316, 335
Jongstra, W., 581
José, Brian, 139n
Juang, B.-H., 610
judgment tasks, 287, 638–45, 660n, 662, 669–70
Jun, Sun-Ah, 198, 456, 473, 528
Jurafsky, Daniel, 628, 686
Jusczyk, Peter W., 664, 669, 705, 737

Ka, Omar, 284, 285
Kabak, Baris, 285, 649
Kachin (Jingpho), 87
Kager, René, 146, 149–51, 675, 677, 744
Kahn, Daniel, 20, 174, 196n
Kahnemuyipour, Arsalan, 454, 463, 482n
Kaisse, Ellen M., 370, 404, 677
Kakehi, Kazuhiko, 649
Kalasha, 243
Kam, 200
Kanerva, Jonni, 217
Kapicioglu, Berk, 336
Kaplan, Aaron, 283
Kaplan, Ronald, 602, 605, 608
Karaim, 245
Karajá, 272, 280
Karanga (Kurunga), 25, 34
Karen, Mark, 335
Karo, 118, 136n
Karttunen, Frances, 27–8, 32, 39n, 608, 620

Kashmiri, 127–8
Katada, Fusa, 134n
Katamba, Francis, 222
Kaun, Abigail, 266, 273, 736
Kavitskaya, Darya, 45, 117, 134n, 136n, 325, 431
Kawahara, Shigeto, 46, 418, 456
Kawakami, Sin, 507
Kay, Martin, 605, 608–9
Kaye, Jonathan, 15, 43, 541, 545, 550, 557, 560, 568n, 735
Kazazis, Kostas, 673
Kean, Mary-Louise, 51
Keating, Patricia A., 382, 528
Kegl, Judy, 711
Kehoe, M., 580, 582, 589
Kekchi Jerigonza, 742
Kelantan Malay, 298
Kelepir, Meltem, 731
Kelkar, Ashok R., 113, 153
Keller, F., 621
Keller, Natalie Eraso, 231
Kemps, Rachel, 335
Kennedy, George A., 525, 527
Kenstowicz, Michael, 27, 43, 45, 49, 52, 60, 65n, 67n, 76–7, 78, 114, 153, 229, 238, 320, 358, 675, 678, 679, 744
Kent, Raymond D., 577
Kenyon, John S., 428–9
Kera, 244, 283, 360
Keresan, 189
Kerkhoff, Annemari, 673
Ketrez, Nihan, 686
Keuleers, E., 391
Keune, K., 379, 383
Key, Harold, 144
Keyser, Samuel J., 20–1, 24, 580
Khalkha Mongolian, 130
Khasi, 95, 301–2, 317n
Khmer, 774–5
Khumalo, James, 244
Kijita (Jita), 77–8
Kikongo (Kongo), 500–1, 527
Kikuchi, Hideaki, 528
Kikuria (Kuria), 282
Kim, Hyunsoon, 316
Kim, Yuni, 87, 224
Kimatuumbi (Matumbi), 34, 269, 467
Kimchi, Ruth, 634, 635
Kimenyi, Alexandre, 294
Kimper, Wendell, 741
Kinande, 216, 281–2, 285, 288
Kindaichi, Haruhiko, 507
King, Lisa, 244, 287
Kinga, 233
Kingdon, Roger, 500
Kingston, John, 214, 309–12, 320, 337, 339
Kinyarwanda, 248–9, 288, 290n, 294
Kiowa, 110
Kiparsky, Paul, 14, 15, 33, 40–1, 42–3, 48, 49, 50, 51, 52, 54–5, 59, 62, 64, 65n, 66n, 68, 75, 80, 82, 95, 98, 120, 121, 136n, 174, 189, 304, 317n, 325, 337, 339, 340, 347n–8n, 370, 402, 404, 405, 408, 409, 410, 412, 541, 576, 678, 680
Kirby, James, 312, 315–16
Kirchner, Robert, 269
Kirk, Cecilia, 589
Kirsner, K., 396
Kirundi (Rundi), 219–20
Kisa, 253, 273
Kisseberth, Charles W., 15, 16–17, 27, 43, 45, 52, 53, 60, 65n, 66n, 67n, 206, 273, 290n, 358, 440, 441–2, 444, 445, 449, 458–9, 480n, 482n, 576, 602, 672, 675, 678, 679
Kitagawa, Yoshihisa, 732
Klamath-Modoc, 15, 140n, 153–4, 608
Klatt, Dennis, 600
Klein, Ewen, 29, 612–13, 614
Klima, Edward, 709, 711–12
Knott, Thomas A., 428–9
Kobon, 153
Kochetov, Alexei, 243
Kohlrausch, Armin, 704
Kolami, 294
Kolodzey, Katherine F., 737, 744, 745–6
Kongo, 500–1, 527
Kontra, Miklós, 286–7
Koo, Hahn, 287, 738
Korean, 83, 295, 316, 528–9, 649
  language games, 732–3, 742, 744–5
Korean-English speakers, 488
Kori, Shiro, 518
Koryak, 94
Koskenniemi, Kimmo, 605
Koutsoudas, Andreas, 27, 59

Koya, 120–1, 137n
Kpelle, 207
Kraft, K., 124
Kraft, M., 124
Krakow, Rena A., 323
Krämer, Martin, 253, 272–3, 276–7, 290n
Kratchovil, Paul, 504
Kratzer, Angelika, 454, 463, 482n
Krauss, Michael E., 331
Krishnamurti, Bhadriraju, 325
Kroch, Anthony, 56
Kroeker, Menno H., 203, 205
Krott, Andrea, 390
Kubozono, Haruo, 170
Kuki-Thaadow, 215
Kukuya, 209–10
Kunana *see* Lardil (Kunana)
Kuria (Kikuria), 282
Kurisu, Kazutaka, 78–9
Kuroda, S.-Y., 52
Kurunga (Karanga), 25, 34
Kuryłowicz, Jerzy, 189, 552
Kwakwala, 453

L2 acquisition *see* second language acquisition
labialization, 46–7, 247, 325
laboratory studies, 512
Labov, William, 305, 325, 335, 403–4, 406–7, 409–10, 411, 413, 423, 433n, 434n
Lac Simon, 735
LaCharité, Darlene, 753, 755, 761, 763, 771, 772, 773, 776
Ladd, D. Robert, 198, 365, 369, 427, 440, 453, 457, 486, 487–8, 497, 498, 513–14, 515, 524, 529, 531
LAFS model (Lexical Access from Spectra), 397
Laganaro, Marina, 653, 654
Lakoff, George, 28, 31–2, 608
Laks, Bernard, 180
Lamoureux, Annie-Karine, 767
Lango, 283, 290n
language acquisition *see* acquisition
Language (Bloomfield, 1933), 351
language change
  and opacity, 40–1, 680
  sign languages, 709–10
  and sound patterns, 319–45, 393, 769
language contact, 754–5
language games, 170–1, 722–48
  definition, 749n
  formal description of, 739, 741–7
  and language acquisition, 723–36
langue, definition, 348–9
Laniran, Yetundo O., 214–15
lapse (stress lapse), 142, 148–9, 150–1
Lardil (Kunana)
  apocope, 27, 28, 45, 91–2
  coda consonants, 110
  epenthesis, 62–3
Large, Nathan R., 655
Larson, Gary, 180
Larson, Meredith, 652, 654
laryngeal harmony, 244–5, 260–1
Lasnik, Howard, 603
Lass, Roger, 612
lateral consonants, 243
lateral relationships, 550
Latin, 48, 140n
Law, Sam-po, 523
layering, morphophonological effects, 73–6, 95, 237–8, 702–5
learnability, 40–1, 54–5, 64–5, 415–16, 594–5, 661–88, 689n–90n, 723–36
  constraints, 665–72
  definition, 661–3
  and interference, 56–7
  phonological alternations, 672–85, 690n
  phonotactics, 663–72
  universals, 573, 577, 663, 666, 685–8, 723–4
Learnability Project (Indiana University), 65n
learner simulations, 316, 396–7, 398–400, 416–23, 431–3, 434n, 665–72, 685–8
  machine learning strategies, 388, 389, 594–5, 621, 622–5, 663
learning biases, 342–5, 688, 729
Leary, Adam P., 368
Lebel, Caroline, 763
Leben, William R., 204, 205–6, 207, 209–10, 366
Lee, Michael D., 642
Lee, Yongeun, 655
left-edge strengthening, 470–1
Leftmost Wins Principle, 69–70
Legendre, Geraldine, 416–17

Lehiste, Ilse, 104, 305, 366
Lehn, Walter, 247
Lena, 254, 269, 276
lengthening, 49, 106–7, 116–17, 441–3, 480n
  ordering of, 66n
  and sound change, 325
Leopold, W., 572
Leti, 90, 126
level-skipping, 439, 441–3, 479n, 483n
  definition, 438
Levelt, Claartje, 581, 582, 583, 588, 591, 689n
Levelt, Willem J. M., 377
Levi, Susannah, 272
Levin, Juliette, 566n
Levine, Robert, 29
Lewis, Geoffrey, 69, 84, 96, 97
lexemes, 307–16
  in sign languages, 693–5
Lexical Access from Spectra (LAFS) model, 397
lexical classes, 55, 74, 84–5
  word-level tones, 199–201, 202, 209–11
lexical decision tasks, 633–4, 646–51
Lexical Morphology and Phonology, 75, 80, 82–3, 98, 304
Lexical Phonology, 35, 370, 402, 403, 404–5, 742–4
lexicon
  lexical minimality, 304
  loanword adaptation, 662, 668, 735, 751–78
  modeling of, 329–30, 368, 392–7, 402, 427–33, 434n, 550–1, 576, 588, 611–12, 626–7, 736, 750n
  neighborhood density, 327–8, 344, 347n, 581, 589, 643
  psycholinguistic models, 633–7, 646–55, 657
Lezgian, 317n, 338
Lhasa Tibetan, 130–1
Li, Bing, 255–6, 278
Li, Boya, 523
Li, Zhiqiang, 504, 507–8
Liberman, Mark, 156, 174, 366, 519, 602, 605
licensing conditions, 284, 296–9, 317n, 535, 539, 540, 550, 557–63, 569n–70n, 587, 652
Lieber, Rochelle, 101, 258
Lieberman, Philip, 327, 505
light syllables, 107, 697
Lillo-Martin, Diane, 696–7, 698
Lindblom, Björn, 326–7, 332, 434n, 516, 577
Line Crossing Constraint *see* No-Crossing Constraint (NCC)
Língua do Pê (language game), 739, 742, 744, 746
Linking Constraint, 21
linking rules, 12, 376–7
Liphola, Marcelino M., 39n, 215–16
liquid consonants, 243
Liss, Julie M., 397, 400
listener-based models, 323–6, 327–8, 329–32, 334, 346n
  learning bias, 342–5
listener perceptions *see* perception cues
Lithuanian, 109–10, 135n, 294, 300
Liu, Fang, 514
Lleó, C., 582, 584, 586
loanwords, 418–20, 428–9, 751–78
  adaptation vs. normalization, 662, 668, 735
  cross-linguistic comparisons, 761–78
  suffixes, 100–1
  unpacking, 774–5
Local, John, 379
local ordering, 744–5
locality theory, 259–60, 261
Locke, John L., 577
Loeweke, Eunice, 231
logical variables, 6, 9–10, 38n, 403–4
Lombardi, Linda, 53, 275
Lomongo (Mongo-Nkundu), 45
Londe, Zsuzsa Cziraky, 287, 414
Longacre, Robert E., 211
longitudinal studies, 576–7, 579, 592–3
Loos, Eugene, 245
Lowenstamm, Jean, 545, 550, 563–4, 569n
Lower Sorbian, 144
lowering (tones), 207–8, 209–10, 212, 221–3, 224
Łubowicz, Anna, 46, 80, 81, 84–5
Luce, Paul A., 397, 398, 400, 647, 650, 664, 669
ludlings *see* language games
Luganda (Ganda), 217–18, 222, 224–5, 234–5
Lukoff, Fred, 48

Lund model, 510
Lunt, Horace, 143
Lyman's Law, 420

Maasai, 53, 203–4, 255, 280
MacDonald, George, 200, 231
MacDonald, Georgetta, 200, 231
MacDonald, Lorna, 158–9
MacEachern, Margaret R., 244, 669
Macedonian, 143, 158
machine learning, 388, 389, 415–23, 431–3, 594–5, 621, 622–5, 663, 665–72
Macken, Marlys, 364, 574, 575
MacNeilage, Peter F., 499, 577
MacWhinney, Brian, 85–6, 393, 579, 593
Maddieson, Ian, 225, 232, 237, 525, 526, 528
Maekawa, Kikuo, 515, 516, 528, 532n
Magnuson, James S., 651
Magri, Giorgio, 689n
Mahanta, Shakuntala, 272, 281, 282
Majors, Tivoli, 266, 283, 284
Makonde, 39n, 215–16
Malay, 126, 298
Malayalam, 100
Mam, 70–1
Mamaindé, 217
Manaster-Ramer, Alexis, 313
map tasks, 380
Mapping Rule, 19–20
Marantz, Alec, 650
Maranungku, 142, 145–6, 156–7, 158, 163
Marentette, Paula, 695–6
Margi, 213, 219
markedness, 12–15, 53, 320, 337–9, 342–5, 541, 543, 651, 745–7
  constraints on, 15, 342, 410–11, 451, 455, 468–78, 483n, 484n, 552, 583, 585–6, 668
  and contrast reduction, 298–304, 309
  feeding relationships, 42–3
  harmony systems, 277
Markovian rules, 4, 38n, 193–4, 604, 610–11, 625, 630n
Marr, David, 632–3
Marslen-Wilson, William D., 633
Martin, James H., 628
Martinet, André, 302, 353
Mascaró, Joan, 51, 80
Massaro, Dominic W., 649
Match Theory, 439–40, 448–58, 467–78, 481n–3n
Mathur, Gaurav, 714
Matisoff, James A., 323
matrices, definition, 8
Matteson, Esther, 150
Matthei, E. H., 574–5
Matthews, Peter H., 683
Mattys, Sven L., 397, 400
Matumbi (Kimatuumbi), 34, 269, 467
Maximal Onset Principle, 615
maximal projection, 453
Maximal Utilization Hypothesis, 43
Maximum Entropy Grammar, 419, 420–1, 434n, 671–2
Maximum Likelihood Principle, 671, 676
May, Jean, 231
Maya, 272–3
McCarthy, John J., 21, 22, 24, 30, 31, 38n, 41, 43, 46, 47, 52–3, 57, 58–64, 65n, 71, 82, 89, 91, 92, 93, 95, 101–2, 152, 174, 264, 267, 275, 277, 297, 450, 451, 458, 470, 544, 575, 676, 689n, 727, 734, 742, 744, 745, 750n
McCawley, James D., 2, 11–12, 31, 230, 435
McClelland, James, 622, 633, 634, 647
McCullough, E., 576–7, 586, 592
McCullough, Stephen, 701
McCune, Lorraine, 499
McIntosh, B. J., 585, 594
McKaughan, Howard P., 200
McLennan, Conor T., 397, 398, 400
McLeod, P., 628
McQueen, James, 284, 368, 634, 650
Meier, Richard, 704
Meir, Irit, 712–13
Melinger, Alissa, 235
Mellander, Evan, 159
Melodic Overwriting, 95–6
memory, modeling of, 388–400, 655–6
Mende, 208
Menn, Lise, 85–6, 576
Mennen, Ineke, 514
Menomini (Menominee), 14
mentalism, 3
Menyuk, Paula, 576
Mercer, R., 599

Merchant, Nazarre, 677
mergers, 291, 305–16, 335, 364
Messer, S., 737
Mester, Armin, 72, 73, 100, 261, 418, 420, 440, 446, 455, 456, 467, 468, 479n, 545, 732
metaphony, 217, 266, 283
metatheories, 2–3
methodology *see also* computational phonology
  articulatory/acoustic studies, 591–2
  artificial language-learning tasks, 343–4, 391, 728–9, 733, 737–8, 746–7
  behavioral experiments, 636–45
  behaviorist studies, 572
  clustering techniques, 311, 312
  commutation tests, 313
  consultants, 650, 759, 776–7
  corpus studies, 380–8, 579, 582, 588, 592–3, 627–8, 656, 752–5, 757–61, 767, 776–7
  descriptive/pedagogical grammars, 205
  elicitation techniques, 201–2, 286, 380–2, 520–1, 582, 593–4, 652–5, 759–60
    infant studies, 287, 591–4, 664
    story-telling, 378–9
  field work, 201–2, 311, 313, 776–7
    diary studies, 572, 574
  formal grammars, 34–5
  "generate and test" method, 611–12
  infant studies, 287, 591–4, 669, 695–6, 728
  instrumental studies, 353, 358–61, 493–5, 520, 650, 660n
    harmony systems, 287
  judgment tasks, 287, 638–45, 660n, 662, 669–70
  laboratory studies, 512
  language games, 170–1, 722–48
  lexical decision tasks, 633–4, 646–51
    nonce tasks, 287, 638–45, 660n, 662, 664, 684–5, 686, 728
    sub-word tasks, 647–8, 651
  longitudinal studies, 576–7, 579, 592–3
  map tasks, 380
  memory tasks, 655–6
  minimal pair test, 353
  reliability, 313, 350–4, 360, 368–71, 372n, 382–4, 428–9, 523–4, 575, 640–5, 722
  repetition tasks, 397
  similarity metric, 262
  statistical analysis, 384–8
  transcription, 200–1, 311, 313, 350–2, 353–4, 362, 365, 382–4, 494, 498–502, 575, 593
  word segmentation tasks, 705–8
Metrical Phonology, 48, 174–5, 456, 487–91, 510, 518–19, 520
  tones, 233, 235–6
metrical stress theory, 145–56, 157–63, 366, 456, 517, 527–30
  acquisition models, 162–3, 574–5, 585–7
  constraints, 146–51, 155–6, 470–2
  and harmony systems, 267–8, 283–4
Meurers, Detmar, 29
Michael, Lev, 114–16, 136n
Michaud, Alexis, 239
Mielke, Jeff, 675, 679
Miller, Corey, 335
Miller, D. Gary, 15–16
Miller, George A., 31
Miller, Wick, 159
MINERVA (computer model), 396–7
mini sound changes, 324, 325
Minimal Generalization Learning, 391
minimal pair test, 353
minimalist theories, 440, 454, 544, 583–5, 681, 683
Minimum Sonority Difference, 177
Minnan (Xiamen) *see* Xiamen (Minnan)
minority languages, 378–9
Mintz, Toby, 287
Mirman, Daniel, 633, 647
Miura, Ichiro, 518
mixed-effects models, 386–8
Mixtec, 210–11, 213, 225
  Acatlán Mixtec, 213
  Peñoles Mixtec, 210–11, 225
Miyata, Yoshiro, 416–17
Mòbà Yoruba, 247
mobile affixes, 87
Möbius, Bernd, 519–20
modality effects (sign languages), 692, 701–8
modeling
  acquisition models, 322–6, 329, 334–5, 338–9, 347n, 369–70, 581, 664–88
    stress, 162–3, 339, 347n–8n
  ambiguity, 605–11, 615, 641–4, 649–51
  Analogical Model, 390

articulation, 327, 332, 395–7, 434n, 503, 504–5, 516
computational methods, 600–2, 605–7, 614–21, 623–5
connectionist models, 622–5, 630n
constituency-based models, 169–71, 178–80, 181–3, 194, 195n, 366, 437–78, 480n–4n, 534
context-free models, 614–16f
derived sonority models, 180–3
episodic models, 395–7, 627
exemplar-based models, 333–6, 347n, 361, 368, 388–400, 627
feature-geometric models, 268–9, 297–8, 320, 544, 581
finite state automata, 184–7, 603–11, 614–15, 618–21, 629n
flat structure models, 174
grid-based models, 156–62
head–dependent relationships, 533–66
inheritance hierarchies, 72, 73
inter-speaker variation, 516
lexicon, 329–30, 368, 392–7, 402
Markovian vs. non-Markovian models, 4, 38n, 193–4, 604, 610–11, 625, 630n
memory, 388–400, 655–6
mixed-effects models, 386–8
multiple-level models, 402–14, 427–33, 607, 626, 634–6, 674–8
n-gram models, 669, 689n
P >> M model, 89–90
phones, 350–8
phonetics-phonology interface, 308–16, 320, 323–45, 361–71
phrase structure rules, 178–9
probabilistic theories, 408–27, 641–4, 649–51
prosodic features, 206–13, 487–530
prototype models, 335, 347n
psycholinguistic models, 632–59, 704–8
quantity, 104–8
speech perception, 646–51, 657–9, 660n
speech production, 651–5, 657–9, 660n
syntax-phonology interface, 435–78
usage-based model(s), 328–30
variation, 327, 388–400, 401–33, 434n, 516, 519, 605–11, 615, 641–4, 649–51, 652, 669–72, 684–5
Mohanan, K. P., 3, 57, 66n, 68, 75, 82, 98, 100, 304, 370
Mohawk, 608
Mokilese, 93
Mongo-Nukundu (Lomongo), 45
Mongolian, 130
monolinguals, role in loanword adaptation, 755–7
monotonicity, 773–6
Montreal French, 763–4, 765
Moon, Seung-Jae, 327, 749n
morae, 107–33
  definition, 135n
  and foot structure, 151–2, 489–91
  geminate consonants, 105–6, 119–27, 129, 134n, 137n–40n
  in language acquisition, 584
  minimal sonority threshold, 109–11, 135n
  moraic mismatches, 129–32, 140n
  and syllable structure, 105–18, 170
    trimoraic syllables, 112–16, 120–2
  and tones, 233, 234–5
moraic phonology, 103, 105–8, 119–27, 135n–40n
Moravcsik, Edith, 93
Mordwin, 113–14
Morelli, Frida, 182
Morén, Bruce, 3, 128, 139n
Moreton, Elliott, 342–4, 733–4, 747
Morgan, James L., 578, 594
Morpheme Structure Constraints (MSCs), 16, 32, 244, 734–6, 738, 748
morphemes
  acquisition, 585–7, 675–85
  ordering, 87–90
  storage of, 393, 427
Morphologically Conditioned Phonology, 97–9
morphology-phonology interaction, 2, 27, 68–102, 402, 608–9, 702–8
  gender marking, 199–200
  harmony systems *see* harmony systems
  in language acquisition, 585–7, 675–85
  layering effects, 73–6, 95, 702–5
  prosodic features, 90–7
    tones, 199–201, 203–4, 209–10, 215–20, 224, 229, 238, 521–4
  sandhi processes, 217–18, 225–9, 238, 377–9, 525–7

morphology-phonology interaction (*Cont'd*)
sign languages, 702–8, 712–17, 721n
syllable structure, 189, 569n
t/d deletion, 405, 406–7
word formation, 83–90, 99–102
word storage, 393, 427
morphomes, definition, 196n
Morrisette, M. L., 587
Mortensen, David, 87
mother-daughter node relations, 31, 470, 546–7
mother talk, 499, 584–5, 588, 701
Motwani, R., 605
Move dialect (Yagaria), 227
movement (sign languages), 693–701, 707–8, 712–17, 720n
Mowrey, Richard, 328
Mpaayci, J. Tompo Ole, 53, 203–4, 280
Mpakwithi, 249
Mpiranya, Fidèle, 249
MSCs *see* Morpheme Structure Constraints (MSCs)
Muller, Jen, 125
Müller, Karin, 615
multimodal corpora, 382
Munro, Pamela, 159
Munson, Benjamin, 327
Murray, R., 555
Muskogee, 94
Mutaka, Ngessimo, 285
mutual bleeding, 62–4, 66n
mutual information, definition, 171
Mwiini (ChiMwiini), 458–60
Myers, Scott, 294, 481n–2n

n-gram models, 669, 689n
N-stratality, definition, 35
Nair, Rami, 733
Nambikuára, 203, 205
Name Game, 731–2
Nanti, 114–16, 128
Nara, 234–5
nasalization, 363
blocking, 245–7, 258, 270–1
"nasal leak", 300
nasals
assimilation, 9–10, 12, 174, 295, 323, 326–7, 404–5, 407, 488
frequency effects, 395
harmony systems, 243, 245–7, 258, 261–2, 266–7, 270–1, 277, 281, 283–4, 289n
nativist theory (acquisition), 573, 577, 663, 666, 672–85, 723–4
nativization processes, 764–71
in sign languages, 696–7
Natural Generative Phonology, 37n
Natural Phonology, 23, 173–4, 357–8
naturalness, and language games, 745–7
Naturalness Condition, 19, 358
Nawuri, 272, 281
Nazzi, T., 594
Ndebele, 215
near mergers, 305–16, 335, 364
negativity, 10–11, 36–7, 203
definition, 5
neighborhood density (lexicon), 327–8, 344, 347n, 581, 589, 643
Nespor, Marina, 100, 438–9, 440, 457–8, 487, 698, 700
Neubarth, Friedrich, 568n
neural activity, modeling of, 36, 635, 704–5
neural networks, 622, 634–5
neutrality, in theory selection, 32–3
neutralization, 39n, 79–80, 294–6, 298, 337–8
in codas, 190, 337–8
in Declarative Phonology, 29–30
definition, 291, 317n
incomplete neutralization, 306–16, 335, 364
modeling of, 321–3, 538, 735–6, 746
Revised Alternation Condition, 14–15, 66n
Nevins, Andrew, 276, 278, 686, 725, 728, 729–30, 731, 739, 740, 742, 744
New World Spanish, 190
New York City English, 305
New Zealand English, 376–7, 381–2, 384–6, 387–8
Newman, Paul, 15, 74, 97–8, 124, 125, 133, 139n
Newman, Rochelle, 650
Newman, Stanley, 52, 557
Newmeyer, Frederick, 387
Newton, B. E., 65n

Ngalakgan, 111–12, 293
Ngambambo, 237–8
Nganasan, 160–2
Nguni Bantu languages, 215
Ní Chiosáin, Máire, 111–12, 271, 274, 302, 304
Nicaraguan Sign Language (NSL), 711, 712
Nichols, Johanna, 69
Nilotic languages, 242
Nimboran, 299
Nishimura, Kohei, 418
No-Crossing Constraint (NCC), 258, 259, 277
Noisy Harmonic Grammar, 416–20, 421–3, 426–7, 431–2
Noll, Craig, 27
non-automatic rules, definition, 54–5
non-derived environment blocking, 80–1
Non-Identity Condition, 39n
non-linear phonology, 18–22, 24–5
non-local vowel-consonant harmony, 249–50, 277–8
non-Markovian rules, 4, 38n
nonce tasks, 287, 638–45, 660n, 662, 664, 684–5, 686, 728
nonlinear phonology, 18–22
Nooteboom, Sieb, 495
Nootka, 46–7
Norman, K. A., 398
Norris, Dennis, 634
Norwegian, language games, 726, 740
notation, 6, 8–12, 13–14, 21, 24, 33–4, 38n, 39n, 66n
  boundary marking, 174
  constituency-based models, 169–71
  evaluation of, 32–3
  head–dependent relationships, 535–7
  null symbols, 186–7
  orthography, 205, 307–8, 317n, 372n, 504, 760, 764–71
  phonetic transcription, 311, 313, 350–2, 353–4, 362, 365, 382–4, 494, 498–9, 575, 593
  representation of quantity, 104–8
  tones, 200–1, 376, 494, 498–502
    X-JToBI tagging, 506–7, 508, 515, 516
  tree graphs, 487, 491–2
  variable rules, 403–4
Noyer, Rolf, 87, 479n, 776
nucleus, 170–1, 184–7, 536–7, 546, 566n
  definition, 134n, 167
Nupe, 96, 200
Nuxalk (Bella Coola), 557

Obligatory Contour Principle (OCP), 24, 26, 30, 209–10, 211, 481n–2n, 650, 745
  definition, 22
obligatory rules, 230–2
Obukuitai, 232
Odden, David, 14, 22, 24, 25, 34, 39n, 136n, 212, 243, 249, 250, 259, 269, 282–4, 467
O'Dell, Michael, 39n, 306, 335
Ohala, John J., 323–6, 327, 328–9, 343, 346n, 503, 505, 723
Öhman, Sven E. G., 520
Ojibwa, 15
O'Keefe, Michael, 52
Olawsky, Knut J., 226–7
Old Quebec French, 753–4, 759, 770–1
Oller, D. K., 576
Onishi, Kristine H., 738
online checking, 391
Onset First Principle, 21
onsets, 116–18, 134n, 136n
  constraints, 20–1, 30–1, 34, 178–9, 188–90, 191–2, 344, 514, 615, 730
  definition, 134n, 166–7
  geminate consonants, 125–7, 139n
  modeling of, 170–1, 181, 184–6, 528–9, 537, 545–8, 552–5, 556–7, 558–63, 566n, 567n
ONZE corpus, 376–7, 381–2
O'odham (Tohono O'odham), 81, 97, 140n, 464–7
Oostdijk, N., 375, 379, 381
Oostendorp, Marc van, 188, 308–9, 426, 552
opacity, 40–65, 369–70, 744 *see also* blocking
  definition, 40–1, 43, 64
  falsification, 41, 64–5
  morphological opacity, 79, 84–5, 678–9
  phonological alternations, 674–8
Optimal Domains Theory, 206

Optimality Theory (OT), 30–1, 35, 54, 80–1, 336, 626, 723–4, 734–5
affix ordering, 87–8
contrast reduction, 308–9, 317n
cophonology vs. constraint indexation, 71–3, 75–6, 82–3
Duke of York derivations, 47
and government phonology, 565
harmony systems, 262–3, 264, 268–9, 270, 271, 278
interface theory, 448–50
and language acquisition, 579–80, 664–7
limitations of, 38n, 59, 194, 354, 370–1, 732, 741–7, 750n
Noisy Harmonic Grammar, 416–20, 421–3, 431–2
Obligatory Contour Principle (OCP), 22, 24, 26, 30, 209–10, 481n–2n, 650, 745
optimal paradigms, 76–80
Partially Ordered Constraints Theory, 408–14, 425–7, 434n
Phonological Readjustment, 89–90
prosodic templates, 91–4
scope of inquiry, 341–2
and sound change, 320, 322, 327, 337, 341
Stochastic Optimality Theory, 391, 414–16, 417, 421–3, 426, 434n, 670–1
stratal theories, 75–6, 98–9
stress, 128, 155–6, 157–8
syllables, 181, 182, 194, 196n, 616–21
underspecification, 20, 541, 587, 611–12
variation, 407–8, 425–7
optional rules, 56–7, 285
ordering, 11, 16, 41–64, 176–8, 539–40
classification of, 43, 66n
constraints, 181, 270–1, 277, 298–9, 301–2, 367, 408–23, 425–7, 434n, 448–50, 466–7, 469, 481n–2n, 565, 583, 591, 594–5, 617–18, 626, 665–72, 685, 689n–90n, 723–4, 726–7, 734
definition, 5, 41–3
in language games, 723–4, 726–7, 734, 739–45
morphology, 74–6, 87–90
and simplicity, 32–3
Ordering Hypothesis (OH), 38n
O'Reilly, R. C., 398
Orgun, Cemil Orhan, 74, 75, 81, 82, 85, 89
orientation changes (sign language), 695, 715–17, 721n
orthography, 307–8, 317n, 372n, 760, 764–71 *see also* International Phonetic Alphabet (IPA)
and tones, 205, 504
Osage (Wazhazhe), 142–3, 146
Osaka Japanese, 518
Ostendorf, Mari, 376
OSTIA algorithm, 686
OT *see* Optimality Theory (OT)
Ota, Mitsuhiko, 583, 584, 585
Otanes, Fe T., 85
overapplication, 43, 58–64, 94–5

P >> M model, 89–90
Padgett, Jaye, 265, 268–9, 271, 274, 302, 304
Pagliuca, William, 328
Pak, Marjorie, 479n
palatalization, 250
Pali, 296–7
Pallier, Christophe, 649
Palmer, Frank Robert, 250
Palmer, Stephen E., 634, 635
Pame, 231
Paolillo, John C., 423, 425
Papoušek, Hanus, 499
Papoušek, Mechtild, 499
paradigm contrast, 78–80
paradigm, definition, 318n
paradigm uniformity, 76–8, 308, 680, 687–8
Paradis, Carole, 26–7, 33, 121, 254, 753, 755, 761, 763, 771, 772, 773–4, 776
parallel constraints, 622–3 *see also* harmony systems
parallel derivations hypothesis, 54
parameters, definition, 23
parametric rules, 23–5, 265–6
parasitic harmony systems, definition, 273
Päri, 242
Parkinson, Frederick, 253, 269
parole, definition, 348–9
Partially Ordered Constraints Theory, 408–14, 425–7, 434n
Particle Phonology, 541
Particle Shift Rule, 5–6
Paster, Mary, 83, 84, 88, 224, 272
Pater, Joe, 73, 418, 419, 428, 581, 583, 667, 671, 676, 685

Patin, Cédric, 239
Pattani Malay, 126
pattern extension (tone sandhi) processes, 525, 526
Paul, Hermann, 323
Paulian, Christiane, 209–10
Pearce, Mary, 244, 283, 360
pedagogical grammars, 205
Peirce, Charles Sanders, 709
Peperkamp, Sharon, 747
perception cues, 594
  first-language acquisition, 364, 576, 577, 588, 592
  harmony systems, 266–7, 274
  misinterpretation of, 323–6, 643–4, 647–51, 660n
  modeling of, 410–12, 622, 633–7
  segmental level, 705–8
  and sound change, 322–6, 327–8, 329, 330–2, 342–5, 347n
    contrast reduction, 299–316, 321–3
  storage of, 395–7, 398–9
  tones, 503, 512–13, 514–15, 520–1, 522, 525
  visual vs. auditory processing, 704–5
  voicing, 311, 321–3
Perceptron convergence proof, 419
perfective aspect, 97
performance, 651–5, 657–9, 660n
Perlmutter, David M., 695, 713, 714
Peters, P. S., 602, 604
Peterson, Gordon, 104, 305
Petitto, Laura A., 695–6
pharyngealization, 247–8, 289n
Philadelphia English, 407
Philippson, Gérard, 216, 223
Phillips, Betty S., 325, 430
PHON (computer software), 593
phonemes, definition, 193, 304–5, 310, 350, 353–4, 355, 370–1
phones, modeling of, 350–8, 372n
phonetic encoding, 651–5
phonetic transcription, 311, 313, 350–2, 353–4, 362, 365, 368, 382–4, 494, 498–9, 575, 593
phonetics, definition, 348–9, 371n–2n
phonetics-phonology interface, 2, 348–71
  contrast reduction, 299–304, 307–16
  covert contrast, 309–16, 364, 575–6, 592
  definition, 348–9
  diachronic sound patterns, 319–45
  harmony systems, 266–7
  modeling of, 308–16, 320, 323–45, 361–71
Phonetics (Pike, 1943), 351–2
phonological alternations, acquisition of, 672–85, 690n
phonological contrast *see* contrast, definition
Phonological Phrase (φ), 443–8, 449–50, 453–4, 455–6, 457–68, 469–78
  definition, 285, 437
  downstep, 473–8
  in sign languages, 697–700, 721n
Phonological Readjustment hypothesis, 89–90
phonological representation, definition, 29
phonological theories *see* theories
phonological variation *see* variation
phonologization processes, 313–16, 323–45
phonology, definition, 348–9, 371n–2n
phonology-morphology interaction *see* morphology-phonology interaction
phonotactics, 183–7, 539–40, 638–59, 705–8, 737–8
  acquisition, 663–72
  definition, 195n
  modeling of, 2, 551–7, 600–2, 605–11, 624–5, 638–45
  phonetic encoding, 651–5
  prelexical processing, 646–51, 657
phrase boundary marking, 155–6, 224, 457–68, 469–78, 484n, 508, 509, 698, 699f, 700–1
phrase structure rules, 178–9, 612, 614
Phuthi (Swati), 91
Picard, 139n
Picard, Marc, 196n
Pickering, M. J., 386
pidginization, sign languages, 711, 712
Pierrehumbert, Janet, 29, 79, 155, 156, 262, 328, 335, 344, 347n, 365, 376, 398, 427, 432, 456, 460, 487, 498, 503, 507, 509, 513, 515, 516, 518, 519, 524, 527, 610, 628, 658, 733
Pig Latin (language game), 725–6, 729, 740, 748
Piggott, Glyne, 15, 247, 259–60, 271, 283
Pike, Eunice V., 168, 213

Pike, Kenneth L., 168, 171–2, 183, 199, 351–2, 362, 503–4
Pinker, Steven, 622
Pinnow, Eleni, 395
Pintupi, 142, 146, 150
Pirahã, 118, 136n
Piro, 150
Pisoni, David B., 655
pitch accent, 230–6, 483n–4n, 509
  and constituent structure, 435, 436, 460–1, 470
  definition, 230, 503–4
  downstep, 460–1, 473–8
  Leftmost Wins Principle, 69–70
  and stress, 155–6, 462–3
  tonal space, 475, 515–21
pitch track, definition, 493
Pitt, Mark A., 375–6, 381, 383, 650
place assimilation, 9–10, 12, 295, 297, 326–7, 331–2, 375–6, 404–5, 407, 488 *see also* harmony systems
place names, 74
Plag, Ingo, 390
Plains Cree, 378–9
plateauing, 217–19, 220, 224–5
Plauché, Madelaine, 347n
Plénat, Marc, 726–7
Ploch, Steven, 321
Plug, L., 379
plurals, constraints on, 85–6
Pluymaekers, M., 377
POC theory *see* Partially Ordered Constraints Theory
Pöchtrager, Markus Alexander, 564
point of articulation, 174, 194
polarity, tones, 219–20
Polgárdi, Krisztina, 552, 565
Poliquin, Gabriel, 679–80
Polish, 58–9, 84–5, 679
  syllable structure, 554, 557
Polysp (Polysystemic Speech Perception), 398
Ponapean, 135n
Poplack, Shana, 777
Port, Robert F., 39n, 306, 307, 335, 368
Portuguese, 584, 739, 742
Poser, William J., 69–70, 71–2, 85, 258
possessives, 85–6, 229
possibility, definition, 2–3
Post, Emil, 4, 597
post-lexical rules, 402, 404–5
post-velar harmony *see* emphasis harmony
Postal, Paul M., 19, 354
Potts, Christopher, 452
Praat (computer software), 416, 417, 419–20, 493, 593
pragmatics, 379, 523–4
Prague School, 295–6, 355, 573
pre-verbal melodies (intonation), 499
Precoda, Kristin, 237
prefixes, 219–20
prelexical processing, 633, 646–51, 657
Preliminaries to Speech Analysis (JFH, 1952), 355–6
Preservation Principle, 763, 774–6
Prieto, Pilar, 585
Prince, Alan, 21, 31, 65n, 66n, 71, 82, 89, 91, 92, 93, 95, 101–2, 135n, 156, 157, 159, 174, 236, 264, 298–9, 320, 366, 408, 450, 451, 458, 470, 579, 580, 602, 605, 617, 622, 672, 675, 676, 734, 745
Principles and Parameters theory, 26–7, 33, 35, 36, 602
Principles of Phonology (Trubetzkoy, 1939), 348–9, 352, 355, 371n–2n
Prinz, M., 582
probabilistic theories, 403–4, 414–27, 615, 639–40, 641–4, 649–51 *see also* variation
progressive aspect, 93
projection, 535–9
pronunciation variation *see* variation
Proper Government Principle, 550, 557
prosodic features *see* prosody
Prosodic Homogeneity, 100
Prosodic Licensing, 26, 587
prosodic morphology, 90–7, 100–2
Prosodic Phonology, 378
prosodic phonology frameworks, 500–21
Prosodic Word (ω), 453–6, 468–9, 470, 478, 490–1, 517
  acquisition, 583–5, 586
  definition, 437
  in sign languages, 697–700
prosody *see also* intonation; syntax-phonology interface; tones
  acquisition, 583–7, 669–70
  category types, 437

definition, 366, 372n–3n, 486, 487–8, 530–2, 532n
formal description of, 18–20, 26, 33–4, 38n, 198, 207, 365–7, 436–40, 450–78, 487–530, 609–11
sign languages, 697–701
tree graphs, 487, 491–2, 610
and language acquisition, 574–5
long-distance effects, 224–30
morphophonology, 69–71, 72–3
sign languages, 692–701, 720n–1n
prototype models, 335, 347n
Prunet, Jean-François, 761, 773–4
Przezdziecki, Marek, 284, 324
psychological capacities, definition, 634
psychological realism, 631–59
definition, 632–7
and language acquisition, 576–8
and linguistic theories, 645–58
and research methods, 636–45, 657–9, 705–8
Puerto Rican Spanish, 307
Puinave, 232
Pulaar, 88, 254–5, 270, 280
Pulgram, Ernst, 164, 172–3, 194, 195n, 196n
Pulleyblank, Douglas, 20, 21, 24–5, 53, 213, 216, 219, 247, 254, 255, 259, 260, 265–6, 270, 276, 280, 281–2, 282–3, 285, 290n, 295, 297, 320, 541
Pullum, Geoffrey K., 47, 274
Püschel, Dirk, 704
Pycha, Anne, 140n, 287, 746–7
Pylkkänen, Liina, 650
Pylyshyn, Z. W., 622

quantifiers, 39n
quantity, 103–33, 134n–40n, 305, 377–8, 458–60, 551–7, 560–1, 568n–9n, 583–4, 708
codas, 105–12, 128–9
constraints, 46, 106–7, 116–17, 441–3, 480n
context dependent weight, 127–9, 135n–6n, 138n, 139n–40n
and contrast reduction, 314–15
definition, 103–4, 139n
formal description of, 104–8, 140n, 441–3
moraic mismatches, 129–32, 140n
onsets, 116–18, 134n
in sign languages, 697
and sound change, 325
and stress, 117–18, 122–3, 128, 130–2, 136n–40n, 142–5, 151–4, 162–3, 428–9
trimoraic syllables, 112–16, 120–2, 142–63, 552, 555
Quebec French, 759, 763–4, 775
loanword adaptation, 753–4, 768–9, 771
stress, 527, 528
Quechua, 144–5, 148, 150
South Conchucos Quechua, 144–5, 148, 150
questionnaires, 286, 287

Rabiner, L., 610
radical CV phonology, 544, 563, 567n
Raimy, Eric, 729, 739, 740, 749n
raising, 46, 51, 58–9, 254
tones, 207–8, 212–13, 221–2
Ramus, Franck, 594
range (tones), 515–21, 529
Rapp, Brenda, 362–3, 654
Rathmann, Christian, 714
Realizational Morphology, 97–9
recursion, 439, 441–3, 452–3, 454–7, 479n, 566, 567n
definition, 438
Recursive Constraint Demotion (RCD), 665–9, 676
reduction, 78, 377–8, 394
in language acquisition, 574–5, 580–5
in sound change, 328
reductionism, 633, 635
redundancy, 16
reduplication, 65n, 93–7, 101–2, 749n
in language games, 729–30, 731–3, 739–41
prosodic templates, 92, 93–4
in sign languages, 697
Reenen, P. Th. van, 430
register variation, 379, 425–7
Reiss, Charles, 3, 11, 22, 36, 39n, 368
relationship rules *see* head–dependent relationships
relative clauses, filtering, 7
Relevancy Condition (RC), 14

reliability, 313, 350–4, 360, 368–71, 372n, 382–4, 428–9, 523–4, 575, 640–5
Ren, Hongmao, 228–9, 525, 526
Rennison, John R., 568n
repair, 6–7, 12–16, 540
Repeated Morph Constraint (RMC), 85–6
repetition tasks, 397
research questions, 2–3, 32–3, 36–7
  blocking, 11
  corpus studies, 387–8
  language games, 723–4
  opacity, 64–5
  phonetics-phonology interface, 319, 341, 358, 362–4, 367–71
  psycholinguistics, 657–9
  sign languages, 706
  syntax-phonology interface, 436, 463, 464, 467–8
  tones, 205–6, 238–9, 509, 514–15, 520–1, 524
  variability, 309, 432–3
researcher bias, 313, 360, 372n, 382–4
Resyllabification Convention, 21
retroflex consonants, 293–4, 300
  harmony systems, 242–3, 248–9, 271, 274, 290n
Revised Alternation Condition (RAC), 14–15, 51, 66n
Reynolds, William, 409
rhinoglottophilia, 323
rhotic consonants, 376–7, 384–6, 581, 756, 774–5
  in harmony systems, 243
rhyme, 170f, 171, 185f, 186–7, 535–7, 545–8, 549–51, 552–3, 555, 556–63
rhythmic stress systems, 142–4, 149–51
Riad, Tomas, 198
Rialland, Annie, 222
Ribeiro, Eduardo, 272, 280
Rice, Keren, 675
Ridouane, Rachid, 505
Rietveld, Toni, 386, 520
Riggle, Jason, 94, 666, 671
Riha, Helena, 504
Rijk, Rudolf de, 46
Ringen, Catherine O., 137n, 286–7, 340
Ritchie, R. W., 602, 604
Ritter, Nancy, 164, 188, 538, 547, 549, 557, 563, 564
Rivard, Adjutor, 759
Rizzi, Luigi, 452
Roach, P. J., 172
Roark, Brian, 584, 588
Robinson, J. J., 534
Roca, Iggy, 673
roots
  definition, 81
  harmony systems, 220, 242–3, 279–81, 284–6
  phonological alternations, 75, 76, 77–8
  prosodic features, 91–2, 220, 224
  root-faithfulness principle, 82, 220
  in sign languages, 693–5
  see also boundaries
Rose, Sharon, 241, 244, 250, 261, 262, 278, 287
Rose, Yvan, 579, 581, 582, 593
Rosenberg, C. R., 622
Rosenbloom, P. C., 597, 599
Rosenthall, Sam, 127–8, 140n
Ross, John Robert, 6–7
Rotokas, 232
rounding (consonants), 333–4
rounding (vowels), 250, 252–3, 259, 265, 266, 268–9, 272, 273, 543
Rowicka, Graznya, 565
Rubach, Jerzy, 553, 554
Rudy, J. W., 398
ruki rule (Sanskrit), 10
rules, 1–25, 31–7, 597–611 *see also* notation
  cyclic rules, 605
  definition, 3–4, 8, 38n
  iteration, 12, 14, 38n
  language games, 725–36, 739–45
  Markovian vs. non-Markovian, 4, 38n, 193–4, 604, 610–11, 625, 630n
  multiple-level systems, 402–14, 427–33, 607, 626, 742–4
  opacity *see* opacity
  optionality, 56–7, 285
  ordering *see* ordering
  parametric rules, 23–5, 265–6
  phrase structure rules, 178–9, 612, 614
  underapplication, 44–58
  underdetermination, 726, 727–36
  variable rules theory, 404–5, 423–7
Rumelhart, David E., 622
Rumsey, Alan, 300

Rundi (Kirundi), 219–20
Runyambo-Haya dialect, 129–30
Russell, K., 378–9
Russian, 79–80, 568n, 629n, 760, 774, 777n
Russo, Tommaso, 710

S-pruning, 6–7
Sabimana, Firmard, 219–20
Saeed, John, 237
Safwa, 504
Sagey, Elizabeth, 21, 259, 261
Salish, 278, 289n
  Lushootseed, 95, 293
  Snchitsu?umshtsn, 249–50
  Spokane-Kalispel-Flathead, 249
Saltzmann, Elliot, 518
Samothraki Greek, 136n
Sample, Ward, 253
San Francisco del Mar Huave, 87
Sanders, Gerald, 27
Sanders, Nathan, 266, 277, 679
sandhi processes, 217–18, 225–9, 238, 376–7, 378–9, 525–7
Sandler, Wendy, 696–7, 698, 700
Sankoff, David, 404, 423
Sanskrit, 10, 248–9, 271
Santa Ana, Otto, 407, 429
Sapir, Edward, 46, 49, 152, 631
Saporta, Sol, 170
Sardinian, 135n
Sasa, Tomomasa, 280
Satta, G., 620
Saussure, Ferdinand de, 167, 355, 573, 638, 658–9
Savà, Graziano, 241
Sawusch, James R., 650
scalar weight, 152–4
scaling (tones), 515–21, 529
Scarborough, Rebecca, 328
Schachter, Paul, 85, 210
Schacter, D. L., 396
Schane, Sanford A., 370, 541
Scheer, Tobias, 549, 550, 556, 563, 564, 568n
Scheibman, Joanna, 378
Schepman, Astrid, 514
Schouten, M. E. H., 430
Schuh, Russell G., 219
Schuppler, B., 381
Schütz, Albert, 159–60, 753
Schütze, Carson, 641
schwa-deletion, 15, 38n, 308, 394–5, 430–1, 558–9, 569n
schwa-reduction, 299
Schweickert, Richard, 655
Scobbie, James M., 29, 37n, 364, 592
Scots (Scottish English), 272
Scottish Gaelic, 246, 270–1
second language acquisition, 56, 369–70, 488
secondary articulation, 245
Seediq, 336
segmental level, 18–22, 24–5, 350–8, 486, 503, 541–5, 597–603
  definition, 349–50, 356, 372n–3n
  perception cues, 705–8
Seidl, Amanda, 747
Sejnowski, T. J., 622
self-destructive feeding, 59–60, 66n
Selkirk, Elisabeth O., 119, 120, 137n, 156, 171, 175, 177, 435–6, 438, 439, 450–1, 453, 454, 456, 458–62, 466, 469, 479n, 482n, 574–5, 585–6
Selkup, 122, 137n
Semitic languages, 344, 555
Seneca, 235–6
Senghas, Ann, 711
Serbian, 216, 515
Sereno, Joan, 335
serialism *see* ordering
Servigliano Italian, 217
Sesotho, 579
Seto (Southeastern Estonian), 121
Sezer, Engin, 74, 105
Shademan, Shabnam, 642–3
Shahin, Kimary, 247, 274, 289n
Shambala, 212, 223
Shannon, C. E., 596, 604
Shared Features Convention, 21
Shattuck-Hufnagel, S., 592
Shaw, Patricia A., 69, 241, 258, 260–1, 289n, 370, 404, 552
Shen, Tong, 453
Shepherd, Michael A., 616–17
Sherer, Timothy, 120, 121
Sherzer, Joel, 735
Shibatani, Masayoshi, 17–18, 38n

shift (tone shift), 215–17, 234–5, 238
Shimron, Joseph, 641
Shin, Eurie, 747
Shinasha (Borno), 221
Shinya, Takahito, 456
shm-reduplication (language game), 729–30, 731, 737–8
Shona, 273, 702
shortening, 15, 49, 66n, 560–1
Shoshone, 159
Shosted, Ryan, 747
Sibe, 250, 278
sibilants, 242, 258–9, 260–1, 262
Sievers, Eduard, 687
sign languages, 691–720
  acquisition, 695–6, 701, 709, 711–12
  iconicity, 692, 709–19
  modality effects, 692, 701–8
  structure, 691–701
silent period (acquisition), 576
Silverman, Daniel, 333–4, 347n
Simoneau, Louise, 527
Simonet, Miquel, 307
simplicity (theories), 32–3
simulations
  agent-based simulations, 333, 336
  learner simulations, 316, 396–7, 398–400, 415–23, 431–3, 434n, 594–5, 663, 665–72, 685–8
simultaneous processing, 702–8
Singapore English, 57–8, 63, 66n, 67n
Singh, Rajendra, 27
Sinological transcription conventions (tone transcription), 499
size constraints, 468–9
Sjölander, Kare, 493
skeletal positions, 539, 545–6, 568n, 707–8
Skinner, B. F., 572
Skoruppa, Katrin, 747
Skousen, Royal, 390, 391
Slavic, 13, 549
Slobin, Dan I., 673
Smiljanić, R., 515
Smith, Eric, 423
Smith, Jennifer L., 327, 667, 733–4
Smith, N. V., 369, 370, 573
Smith, Norval, 233, 258
Smith, R., 398–9
Smoi (language game), 726, 740
Smolensky, Paul, 31, 66n, 71, 269, 298–9, 320, 408, 417, 470, 579, 580, 617, 621, 635, 664–7, 675, 723, 734
Snider, Keith, 209
sociolinguistic variation, 379, 425–7, 754–5, 777
software, 416, 417, 419–20, 423–5, 493, 518, 593, 599–600, 605
  development of, 628
Sohn, H. S., 732–3
Solé, Maria-Josep, 314
Solomon, Nancy Pearl, 327
Somali, 199–200, 233, 237
Sommerstein, Alan H., 18, 38n
Song, Jae Yung, 587, 592
sonorants, Indo-European, 164–5, 167–8
sonority
  definition, 111, 178, 339
  and syllable structure, 109–11, 135n, 166–8, 175–8, 180–3, 193–4, 339, 505, 555, 567n, 580, 617–18, 623–5, 630n, 687
    sign languages, 696–7
Sonority Sequencing Generalization, 177, 182
Sorace, Antonella, 417
Sorbian, 144
Sorsik Sunmakke (language game), 735–6, 742
Sotho, Northern, 442–3
sound change, 319–45
  listeners' role in, 322–6, 330–2, 338–9, 342–5
  modeling of, 313–16, 320, 323–45, 346n, 393
  speakers' role in, 326–30, 331–2
Sound Pattern of English (SPE) theory, 7–16, 172, 354, 356–7, 363–4, 366, 367, 573, 631, 746
Sounding Out (language game), 733–4
Spaka (language game), 737, 744, 745–6
Spanish
  acquisition, 584–5, 586, 588, 590, 673, 682–3
  modeling of, 600–2, 605–7, 614–21, 623–5, 629n, 682–3, 775
  morphology, 76–7
  syllable structure, 171, 173, 190, 307, 334, 568n, 605–7, 614–21, 623–5

SPE (Sound Pattern of English) theory, 7–16, 172, 354, 356–7, 363–4, 366, 367, 573, 631, 746
speakers, role in sound change, 326–30
speech rate, and variation, 404–5
speech recognition systems, 380, 383–4, 599–600, 610–11, 625, 627, 630n
speech synthesis, 600, 627
spell-out processes, 652, 654
Spencer, Andrew, 574
spirantization, 59, 77–8
split ergatives, 341
Spoken Dutch Corpus, 375, 379, 381
spreading, 19–20, 25, 26
  harmony systems, 257, 258–60, 264–6, 270–1, 272, 274–7, 279–80, 281
  sign languages, 698
  tones, 208, 211–13, 214–17, 220, 221–30, 443–8, 449–50, 480n–2n
Sproat, Richard, 100
Sprouse, Ronald, 85, 89, 680
SSCs *see* surface structure constraints
Stallcup, Kenneth, 217
Stampe, David, 23, 574
Stanford Child Phonology Project, 576–7, 578, 726
Stanley, Richard, 13, 16, 32, 38n
star-parenthesis notation, 12, 14
Staroverov, Peter, 45
statistical analysis, 384–8
Stemberger, Joseph P., 170, 286–7, 393, 587, 654
stems *see* roots
Steriade, Donca, 21, 53, 96, 130, 177, 256, 260, 261, 284, 291, 293, 296, 300, 301, 304, 308, 311, 321, 410, 431, 541, 667, 687, 740
Stevens, Kenneth N., 238–9, 310, 599, 627
Stewart, J. M., 52
Stites, Jessica, 580, 589
Stivers, Tanya, 300
Stochastic Optimality Theory, 391, 414–16, 417, 421–3, 426, 434n, 670–1
Stockall, L., 650
Stockholm Swedish, 510–13, 529
Stoel-Gammon, Carol, 575, 580, 587, 589, 592
Stokoe, William, 691
Stoll, R. P., 204
Stone, G., 398
Stong-Jensen, Margaret, 14
stops, 307, 321–3, 575–6
  aerodynamics, 324
story-telling, 378–9
Stratal Optimality Theory, 75–6, 98
Stratificational Grammar, 630n
stress, 390, 435, 463–4, 469–70
  acquisition, 579
  and contrast reduction, 294, 302–4, 315, 339, 347n–8n
  definition, 141, 530
  extrametricality, 128–9, 140n, 235–6
  formal description of, 12, 48, 145–51, 152–63, 366
  and harmony systems, 266, 267–8, 276, 283–4
  and intonational prominence, 154–6, 470–2, 495–7
  metrical stress theory, 145–56, 157–63, 366
  onsets, 117–18
  phrase boundary marking, 155–6
  and pitch accent, 155–6, 462–3
  prominence projection, 136n, 527–30
  and quantity, 104, 117–18, 122–3, 130–2, 136n–40n, 142–5, 151–4, 162–3, 428–9
  and tones, 229–30, 235–6, 471–2, 500, 527–30
  typology of, 133, 142–5
  and vowel quality, 113–16, 336
stress-timed languages, definition, 171–2
Strict Cycle Condition, 51, 80
strict layer hypothesis, 437–9, 457, 458, 461, 482n
stricture, definition, 244–5
Stringfellow, Andrew, 650
Strong Start Effect, 470
Structural Analogy Assumption, 534, 566
structure-building change, 292–3
Structure Preservation Principle, 52
structure-preserving change, 292, 325
structure (sign languages), 691–701
Struijke, Caro, 95
Stump, Gregory T., 683
style differences, modeling of, 425–7
Sub-set Problem (acquisition), 667–8, 684
sub-word tasks, 647–8, 651

subcategorization, 84–5
subglottal pressure, 505
subphonemic phonetic differences, 307–16
substance-free phonology, 368
Sudbury, Andrea, 376–7, 384, 386
Sudo, H., 487
suffixes, 68–73, 76–88, 569n
  doubling, 77–8
  and gemination, 121–2
  root-faithfulness principle, 82
  stratal theories, 99
  and vowel harmony, 215–16, 217, 218–19, 255, 280–1, 286–7, 678–80
Sundanese, 243, 246, 258
Suomi, Kari, 266, 284
superheavy syllables, 113, 153, 552, 555
suppletive allomorphy, 83–5
suprasegmental features *see* prosodic features
surface filters, 603
surface-phonotactic approach, limitations of, 2
surface structure constraints, 17–18, 38n
surface-true counterfeeding, 57–8
surface truth, definition, 57
Svantesson, Jan-Olof, 252
Svec, Bob, 281
Swadesh, Morris, 46, 152
Swahili, 458–60
Swati (Phuthi), 91
Swedish, 120, 510–13, 529
Swick, Joyce, 230–1
Swiss German Sign Language (DSGS), 700, 721n
Switchboard Corpus, 377, 380–1, 387
Sybesma, Rint, 523
syllable-timed languages, definition, 171–2
syllables, 105–8, 164–94, 616–21
  ambisyllabicity, 196n, 548, 740
  and boundary marking, 168–9, 174, 522–3, 600–2, 609, 623–4
  codas *see* codas
  constituent structure hypotheses, 168–71, 178–80, 194, 195n
  and contrast reduction, 293–4, 296–7
  definition, 134n, 165–6, 167, 168, 358, 437, 530
  formal description of, 12, 20–1, 169–71, 174–90, 337–8, 528–9, 535–7, 541–63, 566n, 567n–70n, 605–7, 614–21, 623–5
  and harmony systems, 282–4
  in historical linguistics, 164–5
  in language acquisition, 580–5
  language games, 727–34, 737–8, 740
  Metrical Phonology, 174–5
  morphology-phonology interaction, 189
  onsets *see* onsets
  optimization constraints, 83–6
  P >> M model, 89–90
  and phonological alternations, 190–2, 410, 549–51
  sign languages, 695–7
  sonority-based hypotheses, 166–8, 175–8, 180–3, 193–4, 339, 505, 555, 567n, 580, 617–18, 623–5, 630n, 687
    sign languages, 696–7
  and stress, 117–18, 122–3, 130–2, 142–63, 339
  syntagmatic relations, 169–71
  timing, 171–2, 175
  universals, 172–3
  weight, 107–33, 134n–40n, 151–4, 162–3, 489–91, 551–7, 568n–9n
    definition, 107
    in sign languages, 697
    trimoraic syllables, 112–16, 120–2, 152–4, 552, 555
Symmes, David, 499
sympathy theory, 277
synchronization hypothesis, 512–13
syncope, 17, 22, 39n, 527, 528
syntagmatic relations, 169–71
syntax
  constituency-based models, 183, 195n, 437–78
  generative grammar, 5–7, 596–603, 604–5
  head–dependent relationships, 31, 565–6
  recursion, 454–5
syntax-phonology interface, 198, 402, 435–78
  evaluation of theories, 454–68
  and language acquisition, 574–5
  Match Theory, 439–40, 448–58, 467–78, 481n–3n

mismatches, 378, 438–9
strict layer hypothesis, 437–9, 457, 458, 461, 482n
systematic phonetics, 349–71
definition, 349–50
limitations of, 358–64
see also phonetics-phonology interface
Systemic Phonology, 607, 630n

/t~d/ contrast, 293, 306
/t/ deletion, 405, 406–7, 409–13, 421–3, 424–6, 429–30, 431–3, 433n–4n
't Hart, Johan, 495
T-palatalization rule, 15
Tadadjeu, Maurice, 210, 211
Tagalog, 85, 89–90
Tagliamonte, Sali, 423
Tang, Gladys, 700–1
Tanimuca, 231
Tateshi, Koichi, 439, 460
Taub, Sarah, 709, 710
Tauya, 158–9
Tchobanov, Atanas, 180
TDNR template, 88
Tejano English, 407, 422–3, 424–6
templates, 88, 91–3, 575, 615, 658
definition, 91
Tepehua, 243–4
Tera (Cara), 15, 38n
Terbeek, Dale, 306, 312
terminal nodes, definition, 178
ternary rhymes, 546–7
ternary stress systems, 144
Tesar, Bruce, 620, 621, 664–7, 672, 675, 676, 677, 723
Tesfaye, Ashenafi, 221
text-to-speech systems, 600
theories
evaluation of, 1–5, 7–9, 16–18, 31–7, 37n, 38n, 66n, 368–71, 421–3, 454–68, 502–21, 543–5, 563–4
falsification, 27, 28, 39n, 41, 64–5, 340, 375, 541–2, 552
levels of representation, 11, 27–8, 35–6, 37n, 38n, 304–5
neutrality, 32–3
Theory of Constraints and Repair Strategies (TCRS), 26–7, 31
Thibeault, Mélanie, 761, 774
Thomas, Elaine, 221
Thompson, Robin, 709
Thorsen, Nina G., 519
Tiene, 92
Tifrit, Ali, 164
Tigre, 250
TIMBL corpus, 399
timing, 171–2, 175
TIMIT corpus, 380
Tiriki, 224
Tiv, 111, 135n
Tjamoro (Chamorro), 94
Tlachichilco Tepehua, 243–4
Tlale, Maria, 347n
to-behavior, 192–4
Tofa, 286
Tohono O'odham (O'odham), 81, 97, 140n, 464–7
Tokyo Japanese, 232, 474, 484n, 505–9, 515–21
tonal languages, definition, 503–4
tones, 116–17, 129–31, 197–239 *see also* intonation
agreement, 229
anchoring, 500–1, 506–15, 525
and constituent structure, 443–8, 449–50, 458–9, 460–1, 464–7, 480n–1n, 484n
contour tones, 201, 208–10, 215, 300
definition, 199–206, 213–14, 230, 485–6, 530–2, 532n
deletion, 224–5
downstep, 222–3, 224, 376, 460–1, 517–21
effect on ordering of compounds, 87
field work methodology, 201–2
final lowering, 517, 518–19
floating tones, 210–12, 238, 502, 511–12
formal description of, 206–13, 230–6, 237–9, 357, 366–7, 464–7, 483n–4n, 500–27
layering effects, 74–5, 97–8, 237–8
lexical functions, 199, 200–1, 202, 500, 503–4, 507, 511, 514, 521–4
merging, 306, 308
metrical analysis of, 233, 235–6, 527–30
morphological functions, 199–201, 203–4, 209–10, 215–20, 224, 229, 238
notation, 200–1, 376, 494, 498–502
X-JToBI tagging, 506–7, 508, 515, 516
perception cues, 503, 522, 525

tones (*Cont'd*)
phonemicization of, 313–14, 324–5
phrasal anticipation, 224–30
and pitch accent, 230–6
plateauing, 217–19, 220, 224–5
polarity, 219–20
prosodic functions, 524–7
resetting, 460–1
sandhi processes, 217–18, 225–9, 238, 525–7
shift, 215–17, 234–5, 238
spreading, 208, 211–13, 214–17, 220, 221–30, 443–8, 449–50, 480n–2n
and stress, 229–30, 235–6, 471–2, 527–30
tonal space, 515–21, 529
transcription, 200–1, 376, 494, 498–502
tune-to-text mapping, 18–20, 38n, 207–13, 220–30
and voice onset time (VOT), 360
Tonga, 504
tongue root harmony, 254–5, 279
tongue tip-blade, and articulation, 261, 262
Topintzi, Nina, 117, 118, 125, 126, 131–2, 136n, 137n, 139n
Topping, Donald, 94
Totonacan languages, 243–4
Touretzky, David S., 652
toy grammars, 738
TRACE model (lexical processing), 622
Trager, George L., 134n, 306, 312
Traill, A., 293
Tranel, Bernard, 119, 120, 122, 137n
transcription, 311, 313, 350–2, 353–4, 362, 365, 368, 382–4, 575, 593
tones, 200–1, 376, 494, 498–502
transformational grammar, 5–7, 38n, 596–603, 604–5
transparency, 259, 260, 274–7
sign languages, 711–12
see also blocking
transphonologization, definition, 315
tree graphs, 487, 491–2, 610
Treiman, Rebecca, 170, 740
Tremblay, Antoine, 583, 586, 761
triggering conditions, 53–4
Trigo, Loren, 273
trimoraic syllables, 152–4
avoidance of, 112–16, 120–2
Trinidad English, 422f
Trique, 201, 333–4, 347n
trochaic feet, 146, 574–5, 585
trochaic stress systems, 159–60, 470–1
Trommelen, M., 188, 552
Trubetzkoy, Nikolai S., 105, 113, 116, 133, 134n, 135n, 136n, 295, 348–9, 352, 355, 356, 357, 367, 368, 371n–2n, 543, 573, 648
Truckenbrodt, Hubert, 453, 462–8, 471, 479n, 483n
Trukese (Chuukese), 125–6
truncation, 511, 580–5
Tryon, Darrell, 142
Ts'amakko, 241–2
Tsilhquot'in, 331
Tsujimura, Natsuko, 103
Tswana, 336, 347n
Tucanoan languages, 246
Tucker, Archibald, 53, 203–4, 280
tune-to-text mapping, 18–20, 38n, 207–13, 220–30
Tungusic, 255–6
turbidity theory, 308–9
Turkish, 59–60, 105–7, 307
acquisition, 673, 677, 684, 686
morphology, 69, 75, 76, 84, 85, 86, 87–8, 97
reduplication, 96, 731, 749n
stress, 73–6, 81
syllable structure, 555
vowel harmony, 257, 265, 268–9, 272, 673
Tuvan, 251–2, 257, 264–5, 281
Tuyuca, 246, 277
Twi, 210
Twin Sister Convention, 21
Tyler, Stephen, 120
typology, 365, 486, 530–1
and sound change, 336–45

Ubbi Dubbi (language game), 746
Uffmann, Christian, 282
Ullman, J. D., 605
Ulrich, Charles, 159
Umbertide Italian, 272
umlaut (metaphony), 217, 266, 283
underapplication, 43, 44–58, 94–5
underdetermination (language games), 726
underphonologization, 343–4
undershoot, definition, 511–12

underspecification, 20, 541, 587, 611–12
uniformity (paradigm uniformity), 76–8, 308, 680, 687–8
Universal Grammar (UG), 24–5, 36, 340
  and sound change, 320–2, 337–45
universal quantifier, 39n
universality, 5, 36
Universally Determined Rule Application Hypothesis, 59
universals, 24, 82, 172–3, 239, 258, 320–2, 336–45, 621, 733–4
  in language acquisition, 573, 577, 663, 666, 685–8, 723–4
  segmental level, 350–2, 369
  syntax-phonology interface, 439–40, 448–50, 456–7
Unordered Rule Hypothesis, 27
Urarina, 226–7
Urbanczyk, Suzanne, 95
usage-based models, 328–30
Ussishkin, Adam, 79
uvularization, 247–8

Vago, Robert M., 137n
Van-Bik, Kenneth, 202, 231
van den Bosch, Antal, 392, 400
van den Eynde, Karel, 218
Vanelli, Laura, 314–15
VarbRul (computer software), 423–4
variables (logic), 6, 9–10, 38n, 403–4
variation
  conditioned variation, 352–3, 361, 362f
  definition, 401
  frequency effects, 377, 394–7, 429–30
  inter-speaker vs. within-speaker variation, 428–9, 516
  in language acquisition, 576–7, 582, 591, 662–3, 723–4, 725–34
    phonological alternations, 672–85
  in language games, 725–34, 748
  lexical conditioning of, 427–33
  in loanword adaptation, 770–3, 777
  modeling of, 327, 332, 388–400, 401–33, 434n, 516, 519, 605–11, 615, 641–4, 649–51, 652, 669–72, 684–5
    early phonology, 402, 405–14, 427–33
    late phonology, 402, 403–5
    probability distributions, 408–27
    variable rules theory, 404–5, 423–7
  notation, 403–4
  and sound change, 323–45, 754–5
  and speech rate, 404–5
  and subphonemic phonetic differences, 307–16
Vaux, Bert, 65n, 83, 84, 189, 278, 725, 729–30, 731, 740, 744
Vaysman, Olga, 160–2
velar consonants, 250
"velar pumping", 300
Venditti, Jennifer J., 456, 516, 532n
Venneman, Theo, 376, 555
verb agreement, in sign languages, 712–17, 721n
Vergnaud, Jean-Roger, 19–20, 33, 545, 550, 610
Verlan (language game), 726–7
vertical assimilation, definition, 221
Vickers, Douglas, 642
Viemu Picard, 139n
Vieregge, W. H., 382
Vigário, Marina, 584
Vihman, Marilyn M., 499
Vijver, Ruben van de, 172, 196n
Vincent, Nathalie, 768
visibility (head–dependent relationships), 535, 547
visual processing, 704–5
Viterbi algorithm, 611
Vitevitch, Michael S., 641, 642, 647, 650, 651, 653, 654, 669
Vogel, Irene, 100, 285, 438–9, 440, 457–8, 487
voice quality, measuring, 493–4, 520–1, 523–4
voicing, 300–2, 310–12, 315, 321–3
  effect on duration, 104
  harmony systems, 244
  morphophonology, 69
  VOT (voice onset time), 318n, 352–3, 359–60, 575, 755–6
Voorhoeve, C. L., 211, 236, 507
vowel-consonant harmony, 245–50
vowels
  apocope, 27, 28, 45
  articulation, 503
  backness, 251–2, 264–5, 268–9, 290n
  chain shifts, 46, 369–70, 573
  and contrast reduction, 293, 294–5, 302–4, 314–15
  definition, 169

vowels (*Cont'd*)
devoicing, 38n, 39n, 527–8
dissimilation, 219
duration, 78, 104, 113–16, 305, 314–15, 325, 428–9, 458–60, 560–1, 583–4
formal description of, 175, 441–3
harmony systems, 52, 215–17, 218–19, 222–4, 243, 245–56, 257–8, 259, 263–78, 279–81, 289n–90n, 324, 335, 343, 669, 678–80, 690n
definition, 240, 251
umlaut, 217, 266, 283
vowel-consonant harmony, 245–50, 263–8, 270–1, 272–3, 274–8, 281, 282–4, 289n, 295
head–dependent models, 542–3, 549–51
height, 253–4, 269, 270, 276
hiatus resolution, 34, 558–9, 561
in language games, 745–6
nasalization, 363
raising, 254
rounding, 250, 252–3, 259, 265, 266, 268–9, 272, 273, 543
and stress, 113–16, 336
syncope, 17, 22, 39n, 527, 528
tune-to-text mapping, 18–20, 503
Vroomen, Jean, 705

/w/ deletion, 325
Wagner, Michael, 455
Walker, Rachel, 241, 244, 245, 246, 249, 261, 262, 270–1, 274, 276, 277, 281, 284, 287, 288, 290n
Walter, Mary Ann, 744
Wang, William S. Y., 325
Ward, Ida C., 498
Warker, Jill, 652, 738
Warlpiri, 93, 272
Warner, Natasha, 308, 316, 335
Washo, 94, 293
Waterson, Natalie, 574
Watson, Janet, 247, 248
WaveSurfer (computer software), 493
Wazhazhe *see* Osage (Wazhazhe)
Weak Metathesis Condition, 2
Webb, Charlotte, 2
Wedekind, Klaus, 221
Wedel, Andrew, 79, 333, 336
Weenink, David, 593
Weight-by-Position rule, 105–8, 135n, 138n, 139n–40n *see also* quantity
Weight-to-Stress constraint, 128, 157–8
Weissenborn, Jürgen, 578
well-formedness, 6–7, 320, 483n, 602–3, 615, 646–55
definition, 2–3, 38n
and gapping, 260
judgment tasks, 638–45
markedness constraints, 15, 342, 585–6
syllables, 20–1, 177–8, 696
Well-Formedness Condition, 18–19, 20
Wells, Rulon, 171
Welmers, William E., 199, 201, 202, 207
Werle, Adam, 451, 470
Wessels, Jeanine E., 664, 737
West Greenlandic, 15–16
Western Armenian, 83, 84
Western Basque, 46
Wetzels, W. Leo, 232
Wheeler, Deirdre W., 170, 652
White, Katherine S., 594
Whitney, William Dwight, 167, 248–9
Whorf, Benjamin Lee, 184–7
Wiese, Richard, 775
Wilbur, Ronnie B., 95, 700, 710, 714–16f
Wilcox, Phyllis, 710
Wilkinson, Karina, 66n
Willard, Rainbow, 70–1
Willerman, Raquel, 327
Williamson, Kay, 226
Willmond, Catherine, 159
Wilson, Colin, 162–3, 267, 277, 287, 344, 434n, 671, 686–8, 728–9, 747
Wistrand, Kent, 213
Wolf, Matt, 479n
Wolfe, Andrew, 189
Wolof, 284, 285
Wong, Wai-Yi Peggy, 522, 523
word-final codas, 188–90, 191–2, 334, 337–8, 435–6, 582
geminate consonants, 128–9, 139n–40n, 190
word formation, 83–90, 99–102, 344
word-initial onsets, 188, 189–90, 191–2
geminate consonants, 125–7, 139n, 190
word-likeness tasks, 638–45
Word Minimality Hypothesis, 583–5
words, spoken vs. signed languages, 702–8

Wrap XP-plus-Align XP theory, 462–8, 483n
Wright, Richard, 327, 328
writing systems *see* orthography
Wu (Chinese), 524–7
wug tests, 287, 662, 684–5, 686

X-bar theory, 453
X-JToBI tagging, 506–7, 508, 515, 516
Xiamen (Minnan), 238, 460
Xitsonga, 439, 440–50, 469, 470, 480n–1n
XMOD (computer model), 397
!Xóõ (Bushman), 293
Xu, Yi, 504, 508, 514

Yabem, 245
Yaka, 218–19, 243
Yana, 152
Yasinnik, Yelena, 701
Yawelmani (Yowlumni), 16–17, 31, 35, 52, 53, 92, 560–1, 608, 678–9, 680, 690n
yer-alternations, 549, 568n
Yi, Kwangoh, 170
Yidiny, 702
Yip, Moira, 24, 95–6, 208–9, 213, 217, 224, 237, 300, 730
Yokuts, 678–9, 680, 690n
Yoruba, 214–15, 247, 279, 284–5, 295, 297
Younes, Munther, 248
Yu, Alan C. L., 81, 89, 90, 94, 97, 306, 308, 317n, 335, 338, 741
Yucatec Maya, 272–3

Zamuner, Tania S., 581, 589, 673
Zapotec, 203, 293
Zawaydeh, Bushra A., 247
Zec, Draga, 109–11, 113–14, 135n, 182, 216, 233, 296, 439, 468–9
Zee, Eric, 225, 525, 526, 527
Zerbian, Sabine, 442–3
zero representation, 210–13, 224–5, 545, 547, 549–51 *see also* deletion
Zhang, Jie, 300
Zinc, Gaston, 775
Zinza, 136n
Zoll, Cheryl, 72, 73, 94, 95, 97, 101, 102
Zonneveld, Wim, 749n
Zsiga, Elizabeth C., 347n, 373n
Zulu, 215, 244
Zuraw, Kie, 685, 689n
Zuuja-go (language game), 732
Zwicky, Arnold M., 10–11, 683, 735, 749n
Zwirner, Eberhard, 353
Zwirner, Kurt, 353