ROUTLEDGE HANDBOOKS

The Routledge Handbook of Phonological Theory

Edited by S. J. Hannahs and Anna R. K. Bosch

The Routledge Handbook of Phonological Theory

The Routledge Handbook of Phonological Theory provides a comprehensive overview of the major contemporary approaches to phonology. Phonology is frequently defined as the systematic organisation of the sounds of human language. For some, this includes aspects of both the surface phonetics together with systematic structural properties of the sound system; for others, phonology is seen as distinct from, and autonomous from, phonetics. *The Routledge Handbook of Phonological Theory* surveys the differing ways in which phonology is viewed, with a focus on current approaches to phonology. Divided into two parts, this handbook:

- Covers major conceptual frameworks within phonology including: Rule-based Phonology; Optimality Theory; Government Phonology; Dependency Phonology; and connectionist approaches to generative phonology;
- Explores the central issue of the relationship between phonetics and phonology;
- Features 23 chapters written by leading academics from around the world.

The Routledge Handbook of Phonological Theory is an authoritative survey of this key field in linguistics, and is essential reading for students studying phonology.

S.J. Hannahs is a Reader in Linguistics at Newcastle University, UK. He is on the advisory board of the *Linguist List* and annual phonology conferences the *Manchester Phonology Meeting* and the *Old World Conference on Phonology*.

Anna R.K. Bosch is Associate Dean for Undergraduate Programs in the College of Arts and Sciences, and Associate Professor in Linguistics at the University of Kentucky, USA.

Routledge Handbooks in Linguistics

Routledge Handbooks in Linguistics provide overviews of a whole subject area or sub-discipline in linguistics, and survey the state of the discipline including emerging and cutting-edge areas. Edited by leading scholars, these volumes include contributions from key academics from around the world and are essential reading for both advanced undergraduate and postgraduate students.

The Routledge Handbook of Metaphor and Language
Edited by Elena Semino and Zsófia Demjén

The Routledge Handbook of Systemic Functional Linguistics
Edited by Tom Bartlett and Gerard O'Grady

The Routledge Handbook of Handbook of Heritage Language Education
From Innovation to Program Building
Edited by Olga E. Kagan, Maria M. Carreira and Claire Hitchins Chik

The Routledge Handbook of Language and Humor
Edited by Salvatore Attardo

The Routledge Handbook of Language and Dialogue
Edited by Edda Weigand

The Routledge Handbook of Language and Politics
Edited by Ruth Wodak and Bernhard Forchtner

The Routledge Handbook of Language and Media
Edited by Daniel Perrin and Colleen Cotter

The Routledge Handbook of Ecolinguistics
Edited by Alwin F. Fill and Hermine Penz

The Routledge Handbook of Lexicography
Edited by Pedro A. Fuertes-Olivera

The Routledge Handbook of Discourse Processes, Second Edition
Edited by Michael F. Schober, David N. Rapp and M. Anne Britt

Further titles in this series can be found online at www.routledge.com/series/RHIL

The Routledge Handbook of Phonological Theory

Edited by S. J. Hannahs and Anna R. K. Bosch

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2018
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-02581-3 (hbk)
ISBN: 978-1-315-67542-8 (ebk)

Typeset in Times New Roman
by Apex CoVantage, LLC

Contents

Figures

Tables

Contributors

John Alderete teaches linguistics and cognitive science at Simon Fraser University. He uses methods in psycholingumaistics, traditional fieldwork, and theoretical linguistics to probe how complex language systems are learned and used in fluid speech.

Diana Archangeli (PhD 1984, MIT) is Professor of Linguistics at the University of Arizona and an Honorary Professor at the University of Hong Kong. Her research in phonological theory addresses the extent to which phonological systems can be learned without the benefit of an innate linguistic endowment. With Douglas Pulleyblank she is the author of *Grounded Phonology* (MIT Press, 1994).

Iris Berent is Professor of Psychology at Northeastern University. Her research has appeared in journals such as *Science* and the *Proceedings of the National Academy of Sciences*. She is the author of *The Phonological Mind* (Cambridge, 2013).

Ricardo Bermúdez-Otero is based at the University of Manchester. His research focuses on the morphosyntax–phonology and phonology–phonetics interfaces, with particular attention to diachronic issues. He works predominantly on English (especially the Old, Middle, and present-day periods) and on Romance.

Bart de Boer works on the evolution of speech as Professor in Artificial Intelligence at the Vrije Universiteit Brussel. He has published numerous papers on the evolution of speech, both on the evolution of anatomical and behavioural adaptations, and using computer models, experiments, and comparative data.

Anna R. K. Bosch is Associate Dean of Arts & Sciences at the University of Kentucky, and faculty member in Linguistics. Her research interests include phonological theory, Scottish Gaelic phonology, and the history of transcription practices. She has held research fellowships from the American Philosophical Society, the Institute for Advanced Studies in the Humanities at Edinburgh University, and the Thomas J. Watson Foundation.

Diane Brentari is the Mary K. Werkman Professor of Linguistics and Director of the Center for the Study of Gesture, Sign, and Language at the University of Chicago. Her current work focuses on cross-linguistic variation among sign languages, the emergence of new sign languages, and the pro-tactile variety of American Sign Language used by DeafBlind individuals.

Abigail C. Cohn is Professor of Linguistics at Cornell University. Her research focuses on the relationship between phonology and phonetics and is informed by laboratory phonology approaches. She is co-editor of *Oxford Handbook of Laboratory Phonology* and co-organizer of LabPhon 15.

Kearsy Cormier is Reader in Sign Language Linguistics at the Deafness Cognition and Language (DCAL) Research Centre and Department of Linguistics at University College London. She is interested broadly in the linguistic structure of sign languages, especially British Sign Language (BSL), and in visual aspects of language more generally.

Eugeniusz Cyran is Professor of Linguistics and Chair of the Department of Phonology and Phonetics at the John Paul II Catholic University of Lublin. He has authored a book on substantive and formal complexities and licensing in phonology (*Complexity Scales and Licensing in Phonology*, 2010), and a study on the relation between phonology and phonetics based on voicing phenomena in Polish (*Between Phonology and Phonetics Polish Voicing*, 2014), both volumes published by Mouton de Gruyter.

Jordan Fenlon is Assistant Professor in British Sign Language (BSL) at Heriot Watt University in Edinburgh. His research focuses on the linguistics of BSL at the phonological, morphological, and syntactic levels using the British Sign Language Corpus (www.bslcorpus project.org).

Cécile Fougeron is Research Scientist in Experimental Phonetics at CNRS/University of Paris 3 and co-organizer of LabPhon 10. Her research interests include the segmental manifestation of prosodic organization and the phonetic characteristics of speech disorders. She is co-editor of *Oxford Handbook of Laboratory Phonology*.

Stefan A. Frisch is Associate Professor of Communication Sciences and Disorders at the University of South Florida. He studies the processes of speech production, speech perception, and metalinguistic language processing in order to better understand the ways in which language sound structure is organized in the mind/brain.

Nancy Hall received her PhD from the University of Massachusetts Amherst, and is now Associate Professor of Linguistics at California State University Long Beach. Her research currently focuses on the phonetics and phonology of Levantine Arabic.

Michael Hammond is Professor of Linguistics at the University of Arizona. He received his PhD from UCLA in 1984. His areas of interest include statistical phonology, stress and syllabification, English, Welsh, Scottish Gaelic, computational linguistics, psycholinguistics, and poetic meter.

S.J. Hannahs is Reader in Linguistics at Newcastle University. He is author and editor of numerous books and articles on phonological theory, with a focus on Celtic and Romance linguistics. He is also the author of *The Phonology of Welsh* (Oxford, 2013) and the widely used introductory textbook *Introducing Phonetics and Phonology* (with Mike Davenport; 2010, 3rd ed.), and editor of the *Journal of Linguistics*.

Marie K. Huffman is Associate Professor of Linguistics at Stony Brook University. Her research focuses on the acoustic analysis of speech, especially its temporal structure, and the mechanisms underlying speech adaptation. She is co-editor of *Oxford Handbook of Laboratory Phonology*.

Pavel Iosad holds a PhD from the University of Tromsø and is Lecturer in Theoretical Phonology at the University of Edinburgh. He specializes in phonological theory, especially phonological architecture and feature theory, and has a particular interest in the phonology of the Celtic languages.

Martin Krämer is Professor of Linguistics at the University of Tromsø, The Arctic University of Norway. He is the author of *Vowel Harmony and Correspondence Theory* (Mouton de Gruyter), *The Phonology of Italian* (Oxford), and *Underlying Representations* (Cambridge).

Nancy C. Kula is Professor of Linguistics in the Department of Language and Linguistics at the University of Essex. She has worked on a range of issues in phonology including segmental representations, tone and prosody, and phonological interfaces with morphology and syntax. Her empirical focus mainly centres on Bantu languages. She is co-editor of the *Bloomsbury Companion to Phonology* and has published in a wide range of journals.

Neil Myler is Assistant Professor of Linguistics at Boston University. He received his doctorate from New York University in 2014, under the supervision of Alec Marantz. His research interests include morphology, micro-comparative syntax (particularly with respect to English dialects and languages of the Quechua family), argument structure, and the morphosyntax and semantics of possession cross-linguistically. He is the author of *Building and Interpreting Possession Sentences* (MIT Press, 2016).

Heather Newell is Associate Professor at the Université du Québec à Montréal. Her work investigates how morphological phenomena inform theories of phonology, morpho-syntax, and their interface. She is co-editor of *The Structure of Words at the Interfaces* (Newell et al., OUP, 2017) and of the *Canadian Journal of Linguistics*.

Douglas Pulleyblank (PhD MIT) is Professor of Linguistics at the University of British Columbia. His research and publications in phonology have focused on featural properties such as tone and vowel harmony, and his work demonstrates a particular emphasis on African languages, particularly languages of Nigeria. He has collaborated frequently with Diana Archangeli.

Thomas Purnell (PhD Delaware) is Professor of English Language and Linguistics at the University of Wisconsin-Madison. His research interests focus on submodules in the speech chain, particularly tone–stress interactions, the phonetics–phonology interface, and dialectological rules. He is currently serving the American Dialect Society as Editor, *American Speech*.

Michael Ramsammy holds a PhD and MA in Linguistics from the University of Manchester. He is Lecturer in Linguistics at the University of Edinburgh. He specializes in experimental phonetics and theoretical phonology with a focus on the Ibero-Romance languages and English.

Charles Reiss teaches at Concordia University in Montreal. He is co-author of *I-language: An Introduction to Linguistics as Cognitive Science* (2013, with Daniela Isac), *The Phonological Enterprise* (2008, with Mark Hale), and *Phonology: A Formal Introduction* (forthcoming, with Alan Bale).

Tobias Scheer is a senior CNRS researcher (Directeur de Recherche) working at the Université Côte d'Azur in Nice. He has authored a book on Strict CV, *A Lateral Theory of Phonology* (2004); a historically oriented overview guide to the interface of phonology with morpho-syntax, *A Guide to Morphosyntax–Phonology Interface Theories* (2011); and a volume on his own theory of the latter, *Direct Interface* (2012), all at Mouton de Gruyter. He has also published a textbook on syllable structure, *Précis de structure syllabique* (ENS Editions, 2015, in French).

Joseph Paul Stemberger is Professor in the Department of Linguistics at the University of British Columbia. His research focuses on phonological and morphological processing in language production, with a secondary focus on child phonological and morphological development. He finds that linguistic and psychological theories contribute to our knowledge of human language in interestingly different ways.

Paul Tupper is Professor of Mathematics at Simon Fraser University. He uses his expertise in applied mathematics and scientific computation to study systems that arise in phonology, phonetics, and cognitive psychology.

Harry van der Hulst (PhD 1984, University of Leiden) is Full Professor of Linguistics at the University of Connecticut. He has edited over 30 books, and published two monographs and over 170 articles on various aspects of languages. He has been Editor-in-Chief of the international linguistic journal *The Linguistic Review* since 1990.

Jeroen van de Weijer (PhD 1994, University of Leiden) is Full Professor of English Linguistics at Shanghai International Studies University. His interests are phonology, phonetics, morphology, and psycholinguistics and the relations between these linguistic areas. He works on English, Dutch, Chinese, Japanese, and other languages.

Bert Vaux is Reader in Phonology and Morphology at Cambridge University. He is interested in the structure and origins of the phonological component of the language faculty, especially in the realms of psychophonology, historical linguistics, microvariation, and nanovariation. He also enjoys documenting endangered varieties of Armenian, Abkhaz, and English.

1

The study of phonology in the 21st century

Overview and introduction to *The Routledge Handbook of Phonological Theory*

S. J. Hannahs and Anna R. K. Bosch

1.1 Phonological theory and the field of phonology

Many aspects of current phonological theory owe a huge intellectual debt to developments following the publication of the *Sound Pattern of English* (SPE, Chomsky & Halle 1968) in the 1970s, 1980s, and 1990s, often as a reaction to the perceived failings of SPE. Among other innovations that came after or developed out of SPE, the frameworks of Autosegmental Phonology (Goldsmith 1976), Metrical Phonology (Liberman 1975; Liberman & Prince 1977), and Lexical Phonology (Kiparsky 1982, 1985; Mohanan 1982, 1986), along with research establishing suprasegmental phonological structures (Fudge 1969; Selkirk 1980; Nespor & Vogel 1986), all provided a more complex understanding of the principles and objects of the study of phonology, including phonological relations, phonological operations, and phonological structure. Moreover, these developments led to a growing awareness of the autonomy of the phonological module itself, together with acknowledgement of the interconnectedness of phonology with other linguistic components, including syntax, morphology, and semantics. Even approaches advocating a strict separation of linguistic modules recognize that these modules must somehow relate to each other, even if they are essentially autonomous.

Opening a recent issue of the journal *Phonology* at random, or examining a phonological argument or analysis in *Language*, *Journal of Linguistics*, *Natural Language and Linguistic Theory*, or practically any other journal in the field of linguistics, it is striking how often a 'post-SPE' formalism appears, whether exemplified by the use of an autosegmental association, a mention of the lexical vs. postlexical distinction, or a reference to some constituent of phonological structure. Even across differing frameworks and assumptions about the aims, scope, and mechanics of phonological theory, these sorts of representations recur.

It would, however, be wrong to conclude that phonological research over the past fifty years has been unidirectional. Indeed, it seems evident that there is less consensus among practitioners of phonology at present than there was, say, in the 1980s. And of course it would be incorrect to think that there has ever been full agreement as to the proper subject of phonology or the aims of phonological analysis. Indeed, even when SPE was at its most influential there were phonologists rejecting the assumptions of SPE, such as those framing their analyses in terms of Natural Phonology (Donegan & Stampe 1979; Hooper 1976), or those working in a Firthian prosodic framework (e.g. Palmer 1970). Nonetheless, there was a significantly large group of phonologists in the 1970s and 1980s who shared a set of assumptions about the underpinnings of phonological theory, about the aims of phonological analysis, and about where, roughly, to draw the line between phonetics and phonology.

Within current phonology, if we can use that phrase, there are several bases for the lack of consensus across the field at present (though also note Goldsmith's 1992 reflections on fragmentation in linguistic theory in a larger context). Perhaps the most obvious proximate cause of disagreement in generative phonology is the advent of Optimality Theory (OT) (Prince & Smolensky 1993/2002). The analytical framework of OT overturned some long-held fundamental assumptions in generative phonology: namely, the central function of rules to mediate between abstract underlying forms and surface forms; the reliance on phonological rules to derive a single output from an underlying representation; the analytical possibilities afforded by assuming that rules are ordered and that they can interact; the assumption that underlying representations may well be related opaquely to surface forms; and that wordforms in the lexicon may not be directly inferable from surface forms. The popularity of OT among a large group of phonologists, alongside its failure to persuade or engage others, may be said to be responsible for some of the ensuing fragmentation within the field of phonology even among those phonologists who consider themselves generativists, in the sense of Chomsky (1957, 1965).

The development of OT, however, has not been the only factor contributing to the current diversity in phonology, phonological theory, and approaches to phonological analysis. At about the same time as the advent of OT – the late 1980s and early 1990s – other phonologists began questioning some of the assumptions of permissible or desirable abstractness underlying standard approaches to generative grammar (though this debate too can be traced back to the period immediately following publication of SPE, e.g. Schane 1973 vs. Tranel 1981). Within OT this perspective has been particularly associated with the endeavour to ensure that constraints are 'phonetically grounded' (e.g. Pierrehumbert 2000; Archangeli & Pulleyblank 1994; Hayes, Kirchner & Steriade 2004): in other words that the evaluation of output forms be constrained by the phonetic plausibility of the constraints used in that evaluation. Outside OT this concern to tie phonological analysis more closely to phonetic events, to articulation and to language use, is demonstrated in the approaches of Articulatory Phonology (Browman & Goldstein 1986; see Hall, this volume), Exemplar Theory (Bybee 2006; Johnson 2007; see Frisch, this volume), and Usage-Based Phonology (e.g. Bybee 2003).

At least two other recent developments have also led to a certain fragmentation in the pursuit of phonology: sociophonetics and database studies. Sociophonetics has grown out of the Labovian sociolinguistic paradigm, where sociolinguistic variation (such as the variable alternation between [n] and [ŋ] for the English *–ing*, e.g. alternative pronunciations for *singing* such as [sɪŋɪŋ] vs. [sɪŋɪn]) is examined from the perspective of the socio-economic and linguistic contexts associated with the use and variation of pronunciations. For some researchers the issues surrounding this sociolinguist variation, as reflected in the phonetics of differing pronunciations, has become the focal point of investigation, with little investigation into how this type of variation might inform theory and phonological understanding. There are, certainly, exceptions to this – phonologists who are interested in interpreting sociophonetic variation in the light of phonological theory, or who see the value of investigating sociophonetics specifically for how it can inform phonological theory (see Kiparsky

2016) – but they seem to be in the minority. This surface orientation of sociophonetics can lead to a perception that the phonological system itself is of little interest, compared with the study of speech in social context or in use.

The second recent development, database studies, has also led to fragmentation of interest in the field of phonology, and to a reliance on 'big data'. For some practitioners the absence of big data to support a theoretical position reduces the persuasiveness of such a position. While this trend is entirely consistent with recent developments in the sciences in general, it can lead to a refocusing on issues and a readjustment of priorities in phonology: the only worthwhile questions for these linguists are those that we can examine through large corpora of data. As with sociophonetics, this refocusing on the type of data under investigation can lead also to a perception that somehow describing the surface – such as the speech signal via recorded speech – is more important than understanding the underlying system. Again, though, we find a range of perspectives and goals throughout corpus- or data-based studies; there are those working in database studies for the light that these can shed on phonological systems. From a generative perspective, a further difficulty with database studies is that a record of occurring speech cannot by itself explain from surface observation alone why certain structures may be missing accidentally, while others may be impossible. Absence of evidence, however, may simply mean that we haven't found evidence yet: the surface-true record alone, limited as it is, cannot account for the system as a whole.

Despite these caveats, both sociophonetics and database studies do make important connections with phonological research, for instance in Laboratory Phonology (see Chapter 18, this volume) and Exemplar Theory (see Chapter 20, this volume). The observations that are foregrounded by usage-based perspectives are clearly valuable to our wider understanding of speech events, just as observations and measurements from other scientific fields are useful (as in, say, astrophysics). But in phonology just as in astrophysics, there is much more to be understood than the observations themselves; we hope to expand our understanding of phonology beyond mere observation. By not engaging with the systems underlying these observations, we risk missing crucial patterns and interfaces of those systems. Current phonological theory stands at an important crossroads in weighing up the balance between theoretical abstraction and empirical concreteness. While empirical concreteness has contributed enormously to our understanding of what is possible in human language, there is a value also in focused exploration into more abstract elements of language: elements that we cannot see or measure, such as phonological structures, non-surface-true generalizations, and relationships that can only be inferred through theoretical analysis.

In a certain sense, all fields of linguistics can suffer a certain fragmentation brought about by differing perspectives on the field in question, and none of the fragmentation itself is new. The 'functionalist/formalist' split, traceable at least to Saussure (1916), can be seen to play a role in most ways of approaching linguistic study. Yet to take an example from syntax, most 'formalists' or 'functionalists' could agree that apart from mere description there are facts to be accounted for, such as the relationship between a declarative sentence and so-called *wh*-fronting in the related interrogative – how is the prepositional phrase in the sentence *I saw Jane in the pub* related to the interrogative *where* in the question *Where did you see Jane?* This indicates that there is a set of facts to be accounted for, whether one adopts a 'formalist' or a 'functionalist' perspective in accounting for those facts.

When we turn to phonology, however, we are confronted with a challenging question: what counts as a *phonological fact*? How is a phonological fact different from a phonetic fact? *Is* it different from a phonetic fact? Where does phonetics end and phonology begin? Or are they not distinct?

The functionalist/formalist distinction has a particular resonance in phonology, given the position of phonology within the architecture of the language system: phonology is often said to lie at the interface between (mental) grammar and (physical) externalization. While phonetics is typically concerned with physical aspects of the study of speech (externalization), phonology encompasses the connections between the sound system and the grammar of language. In recent years, this distinction has been blurred in various ways, particularly across the domain of phonological theory: usage-based analyses as in the work of Bybee (e.g. 2003, 2010) appeal to facts about speech and language use, such as frequency of a lexical item, to account for observations about language; corpus-based work (Durand, Gut & Kristoffersen 2014) interprets observations about language from the facts of speech; Articulatory Phonology (Browman & Goldstein 1986) suggests that the organization of phonological systems can be attributed to the muscular movements or gestures inherent in the articulation of speech. Even more formal approaches to phonology, such as OT, appeal to the role of phonetic detail, for instance in the arguments for phonetic grounding of optimality constraints (Hayes, Kirchner & Steriade 2004).

These various approaches in recent years have served to blur the distinction between phonetics and phonology, to the extent that some scholars assume that phonology is non-distinct from phonetics (Johnson 2007), while others take the further step that phonology is merely part of the externalization of language and is therefore not part of the grammar at all (Burton-Roberts 2011). Moreover, those scholars who accept that there is a distinction between phonetics and phonology nonetheless draw the demarcation between the two fields of study at different points, e.g. Hale & Reiss (2008), Blaho (2008), and Iosad (2012) arguing that phonology is abstract and thus independent of phonetics (i.e. 'substance-free'), while others such as Hayes, Kirchner & Steriade (2004) and Archangeli & Pulleyblank (1994) argue that phonology is phonetically grounded.

It is against this background of the development over the past thirty years of various ways of doing phonology that this Handbook has been put together. In a real sense, the Handbook is intended to be a snapshot of the current range and breadth of phonological theory, or *theories*. Rather than espousing any particular approach, any single set of assumptions or specific analytical framework, we have sought to encourage practitioners of various approaches to phonology to present overviews of 'their' approach to phonology, thus yielding an overview of the field, written by phonologists of varying stripes or persuasions. While we include a number of chapters written from the generative perspective, this is a reflection of the continuing attraction and importance of generative phonology. We also include various contributions on important issues in phonology from non-generative or even atheoretical perspectives. Contributors were asked to write for a readership of linguists, primarily phonologists, who were not necessarily familiar with the framework, approach, or issues discussed, and to contextualize their approach and make clear how it works and how it relates to other parts of the grammar.

Our hope is that by giving a platform to the many principal approaches to phonology in current practice, and by encouraging our contributors to explain, describe, and exemplify their models to an audience of intelligent readers who are not necessarily familiar with their frameworks and assumptions, we can contribute to the field of research in phonology in some small way. It would be naïve to imagine that simply making explicit the assumptions of a particular model will persuade phonologists of a different stripe to adopt that model. At the same time, by explicitly describing the phonological models in contemporary usage, one may hope that misunderstandings based on misrepresentations will be attenuated.

1.2 Structure of the volume

The Handbook is divided into two parts, the first part dealing with analytical approaches to phonology, the second part looking at a wider range of issues relevant to phonological theory.

Given the diversity and varying positions even on basic assumptions in current phonology, there are few strict lines of demarcation in the field. As a consequence, it can be difficult to structure a volume of this sort to reflect objectively the many threads of investigation, the differences in conceptualization, starting points, and goals, while at the same time trying to achieve coherence of coverage. There will, no doubt, be some readers who will object that their preferred approach has not been given the prominence it merits. In structuring the volume we have decided to foreground major theoretical approaches in Part I: in other words, to present in some detail the most influential analytical frameworks used in current phonology. These include Optimality Theory (Chapters 2–5), Rule-Based Phonology (Chapters 6–8), Government Phonology (Chapters 9–11), Dependency Phonology (Chapter 12), and connectionist approaches to phonology (Chapters 13 and 14).

Part II, in contrast, presents something of a smorgasbord of issues – rather than introducing specific theoretical frameworks, this part of the Handbook examines issues in phonology that to an extent are theory-independent, such as the role of phonetics in phonological analysis (Chapter 15), or that of articulation (Chapter 19); the origins of phonological objects whether as part of Universal Grammar or emergent in the grammar of the speaker (Chapter 17); or the role of analogy in phonology (Chapter 20). Beyond such issues, these chapters also address more general questions of phonological data collection and analysis (Chapter 18), along with questions of how different modalities can be modelled in phonology (Chapter 16). The role of mathematical models in phonology is addressed (Chapters 21 and 22), and the biological origins of phonology itself are explored (Chapter 23).

1.2.1 Part I

Part I begins with three chapters laying out the priorities of Optimality Theory (OT) that survey some of the important empirical and conceptual contributions of this constraint-based model, particularly as regards phonological 'conspiracies'. In Chapter 2, Pavel Iosad argues that OT addresses issues of learnability more effectively than other theories, and benefits from its ability to incorporate explicit predictive mechanisms. In Chapter 3, Martin Krämer provides a detailed outline of the types of constraint interaction that have been proposed by phonologists working in OT, demonstrating that constraints can be organized and interact in a wide variety of ways. This chapter discusses the use of parallel and serial computation, and foreshadows some of the issues examined more closely in Chapter 5 dealing with Stratal Optimality Theory. Michael Ramsammy illustrates in Chapter 4 how the different modules of grammar can be integrated into a coherent whole through constraint-based theories such as OT, with a particular focus on phonological alternations that are conditioned by morphological processes. These related chapters on OT lead into Chapter 5, in which Ricardo Bermúdez-Otero introduces Stratal Phonology, adapting the principles of constraint-based phonology within a cyclic model. Stratal Phonology presupposes that morphology and phonology are distinct grammatical modules, and so favours nonlinear approaches to apparently nonconcatenative exponence.

In Chapter 6, Thomas Purnell presents the first of three chapters on Rule-Based models of phonology, providing a detailed explanation of the representations and formalisms of phonological objects, on the one hand, and functions or rules, on the other. The section on rule formalism provides explanations of the use of disjunctive rule application, feature matrices, and autosegmental features and grounding; he goes on to argue for a representation of segments which incorporates a contrastive feature hierarchy (Dresher 2009). Chapter 7, by Bert Vaux and Neil Myler, surveys some of the main predictive differences between rule- and constraint-based formalisms that have taken centre stage in work on Rule-Based Phonology since 1993. In addition to the familiar cases of conspiracies (argued to favour OT) and opacity, the authors review unnatural processes, morpheme structure constraints, and local (as opposed to global) process interactions. This chapter concludes by suggesting prospects for future exploration of language in non-auditory modalities as a window into more abstract properties of phonological and general computation, and outlines a possible cross-over between rationalist and empiricist phonological perspectives using information theory and/or Bayesian probability. The final chapter on Rule-Based approaches to phonology, Chapter 8 by Heather Newell, provides insight into the ways Rule-Based Phonology can effectively address interfaces between or among various modules. This discussion is centred around questions of modularity and derivation/cyclicity, such as (i) investigating the necessity of both derivational and representational accounts of phonological domains; (ii) detailing how constraint-based theories differ from rule-based theories as to how cycles are triggered, and in which module (the syntax or the phonology) they are controlled; and (iii) questioning whether the effects of constraint reranking across strata can be captured by a theory of structural underspecification.

Three chapters on Government Phonology (GP) provide an overview and outline current issues in the development of this theoretical approach; a fourth conceptually related chapter presents an introduction to Dependency Phonology. In Chapter 9, Tobias Scheer and Nancy C. Kula lay out key assumptions and basic design properties of GP, focusing on the relationship between phonology and phonetics, as well as the way computation works. This chapter includes an outline of element theory, offering a historical overview of the representational theory of elements as contrasted with distinctive features. The basic assumptions and core characteristics of elements are discussed, as well as some of their assumed acoustic correlates. Tobias Scheer and Eugeniusz Cyran outline how GP accounts for syllable structure in Chapter 10. Standard GP does not provide for a syllable node, a coda constituent, or word-final codas. Instead, empty nuclei are given a theoretical status, and syllable structure is constant and present in the lexicon. The hybrid arboreal-lateral system of Standard GP is ultimately revised into a lateral-only approach, known as Strict CV. The latter portion of this chapter details the empirical predictions of this model. The same authors employ GP to address the interface between phonology and morpho-syntax, and between phonology and phonetics, in Chapter 11. This chapter, too, relies on modularity: morpho-syntax, phonology, and phonetics are distinct computational systems that can only communicate through a dictionary-type translation (spell-out). In Chapter 12, Harry van der Hulst and Jeroen van de Weijer introduce Dependency Phonology, which shares certain characteristics with GP, while representing all phonological relations in terms of dependencies. Dependency Phonology uses a unified set of privative elements rather than binary features; these elements generalize over vowels and consonants as the ultimate units of phonological structure, and account for syllabic organization, as well as organization at higher prosodic levels.

John Alderete and Paul Tupper in Chapter 13 present the first of two chapters on connectionist approaches to phonology. While connectionist models are ubiquitous in psycholinguistics, they are less thoroughly developed as generative models of grammar. This chapter surveys the literature of connectionist models which have been developed to address problems central to generative phonology. The focus of this chapter is to explain precisely how these models work, and in particular how they account for locality, gradience, opacity, and learnability in phonology. An understanding of connectionist phonology yields both a deeper understanding of past developments in phonological theory and a glimpse into its future. Joseph Paul Stemberger discusses interfaces in connectionist phonology in Chapter 14. The three variations of connectionist models are discussed: local (lacking a learning component), distributed non-recurrent (based on automatic learning algorithms, with time encoded along an abstract dimension and only feed-forward processing), and distributed recurrent (based on automatic learning algorithms, with both feed-forward and feed-back, and time encoded in an analog fashion). The success and limitations of these connectionist approaches are critically assessed, and close examination is given to each model's particular assumptions about representation.

Thus the contributions to Part I focus primarily on current analytical frameworks in phonology, detailing the assumptions, principles, and mechanics of these frameworks. Part II, by contrast, deals with a somewhat more disparate set of issues affecting phonology and phonological analysis. These include questions such as the role – if any – of phonetics in phonology, the distinction – if any – between phonetics and phonology, issues of analogy vs. rules in phonology, the connection between spoken language and sign language in terms of phonological analysis and the object of that analysis, data collection and replicability in phonology, and, finally, the evolutionary origins of phonology in the human species.

1.2.2 Part II

Substance-Free Phonology is introduced by Charles Reiss in Chapter 15. As Reiss introduces his subject, he makes clear what Substance-Free Phonology is NOT: this chapter proposes a model of Substance-Free Phonology "that makes no reference to wellformedness, repair, contrast, typology, variation, language change, markedness, 'child phonology', faithfulness, constraints, phonotactics, articulatory or acoustic phonetics, or speech perception." One of Reiss' goals is to identify those aspects of phonology that remain when phonetics, i.e. substance, is removed from consideration.

Jordan Fenlon, Kearsy Cormier, and Diane Brentari present a detailed overview of sign language phonology in Chapter 16. This chapter provides a clear contrast to the previous one in focusing indeed on 'substance', but a visual-gestural substance that is no less abstract when viewed analytically. The authors examine the various phonological models that have been proposed to account for sign language phonology, and their differences, as well as how research on sign language phonology has contributed to our understanding of the relationship between sign language and gesture. The chapter concludes with a brief discussion on how the phonology of sign languages can inform current theoretical issues in the field of phonology generally.

In their chapter on emergent phonology, Chapter 17, Diana Archangeli and Douglas Pulleyblank explore the implications of rejecting a strong innatist position. Their goal is to investigate the tension between the assumption of Universal Grammar (a strong innatist position) and a reliance on general cognition to account for grammar (an emergentist position). Their focus is on emergentist grammar relative to phonology. Arguing that phonological systems do not occur in isolation, these authors look primarily at aspects of the interface of phonology with phonetics and with morphology.

The wide-ranging and diverse field of Laboratory Phonology is presented in Chapter 18 by three linguists active in the field, Abigail C. Cohn, Cécile Fougeron, and Marie K. Huffman.

Laboratory Phonology draws on theories and tools from various branches of the sciences to elucidate the linguistic, cognitive, and communicative nature of speech. In this chapter these authors present an overview of Laboratory Phonology, including the historical context for its development and its key questions, methodologies, and critical results, to illustrate how this interdisciplinary approach has helped advance our understanding of the core questions of concern to phonology.

Nancy Hall provides a clear overview of the methods and concerns of Articulatory Phonology in Chapter 19. The central premise of Articulatory Phonology is that the representational units of phonology correspond to speech production events, thus explicitly conflating phonetics and phonology. Whereas most phonological theories assume that speakers mentally represent a word in terms of the features or segments, Articulatory Phonology uses quite a different set of representations: articulatory gestures, and the coordination structure that determines their relative timing. Gestures act both as units of contrast and as units of speech production, essentially erasing the traditional distinction between phonology and phonetics. This chapter reviews Articulatory Phonology analyses of phonological phenomena including categorical and non-categorical alternations, allophony, syllable structure, moraic structure, morphological structure, tone, and intonation.

Stefan A. Frisch introduces the cognitive model known as Exemplar Theory in Chapter 20. Exemplar Theory explores the use of individual memory traces, and thus the experiences of the individual, in the representation of categories, positing that stimuli – including linguistic stimuli – are processed by comparing them to a set of previous experiences stored in memory. This approach recognizes the contributions of individual speaker differences, contextual cues, and other language experiences in the development of phonological categories and interpretation of linguistic stimuli.

In Chapter 21, Iris Berent provides a clear overview of the experimental methods of Algebraic Phonology. Generative linguistic theories attribute productivity to the grammar's computational system, which is endowed with the capacity to encode abstract equivalence classes, to constrain their constituent structure, and to extend generalizations across the board. These capacities are the hallmarks of the algebraic rules. While the algebraic hypothesis has been tacitly held by most phonological theories, its predictions are rarely tested directly. Moreover, the algebraic view has been vigorously contested by exemplar models and phonetically grounded alternatives. This chapter reviews evidence from psychological experiments and computational simulations in order to address two questions: (a) Are phonological generalizations consistent with the algebraic hypothesis? and (b) Does an account of such generalizations require algebraic rules?

Michael Hammond introduces the reader to the role of statistics in phonology in Chapter 22. This chapter examines those domains where quantitative effects are modelled in phonological theory and where specific statistical phonological machinery is proposed. Variable rules, developed initially to accommodate sociolinguistic variation, are viewed from a rule-based perspective. The author then moves on to examine statistical approaches in constraint-based theories of phonology such as OT, as in Golston (1998), proceeding from there to discuss stochastic OT, developed to account for the mapping between phonology and perception/production, where constraint ranking is probabilistic. The chapter concludes with a prospectus for future developments.

The final chapter ends up where phonology begins: Bart de Boer presents an overview of phonology and evolution, in a biological sense, in Chapter 23. This chapter introduces the basic notions of evolution so as to better prepare the reader for a discussion of the evolution of phonology, with special attention to the similarities and the differences between biological and cultural evolution. Additional topics include the evolution of anatomical adaptations to speech, and proposals for neural adaptations. Evidence is drawn partly from the fossil record, which allows us to partially date the emergence of speech, and partly from comparative evidence with other animals. As evolutionary theory has also been applied to understanding historical processes in terms of cultural evolution, the author discusses how evolution can be used to understand phonological change, how tools developed for the study of biology can be adapted to the study of language and speech, and how historical processes related to cultural evolution can be investigated experimentally. The chapter concludes with a discussion of the interaction between biological and cultural evolution, the theoretical challenges of co-evolving systems, and how these issues can be studied experimentally.

Taken together, the analytical frameworks presented in Part I and the exploration of various issues in Part II provide a snapshot of the current state of phonology, phonological analysis, and the continuing debate about the demarcation, if any, between the study of phonetics and that of phonology.

In compiling this Handbook the editors are painfully aware that we have not achieved exhaustive coverage of the field. Nonetheless, by giving a platform to the principal approaches to phonology currently in use, and by engaging our contributors to explain, describe, and exemplify their models to an audience of linguists who are not necessarily familiar with their specific frameworks and assumptions, we can provide a service to the field in some small way. Moreover, by explicitly addressing various areas of debate in phonology, we hope to increase the possibilities for continued debate, discussion, and continued progress in phonology.

References

Archangeli, Diana B. & Douglas G. Pulleyblank. 1994. *Grounded phonology* (Vol. 25). Cambridge, MA: MIT Press.

Blaho, Sylvia. 2008. The syntax of phonology: A radically substance-free approach. PhD dissertation, University of Tromsø.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology*, *3*(1), 219–252.

Burton-Roberts, Noel. 2011. Where and what is phonology? A representational perspective. In Noel Burton-Roberts, Philip Carr & Gerard Docherty, eds, *Phonological knowledge*, 39–66. Oxford: Oxford University Press.

Bybee, Joan. 2003. *Phonology and language use* (Vol. 94). Cambridge: Cambridge University Press.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language*, *82*(4), 711–733.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1957. *Syntactic structures*. s-Gravenhage, Netherlands: Mouton & Co.

Chomsky, Noam. 1965. *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Donegan, Patricia Jane & David Stampe. 1979. The study of natural phonology. In Daniel A. Dinnsen, ed, *Current approaches to phonological theory*, 126–173. Bloomington: Indiana University Press.

Dresher, Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Durand, Jacques, Ulrike Gut & Gjert Kristoffersen. 2014. Introduction. In Jacques Durand, Ulrike Gut & Gjert Kristoffersen, eds, *The Oxford handbook of corpus phonology*, 1–9. Oxford: Oxford University Press.

Fudge, Erik. 1969. Syllables. *Journal of Linguistics*, *5*, 253–286.

Goldsmith, John. 1976. Autosegmental phonology. PhD dissertation, Massachusetts Institute of Technology.

Goldsmith, John. 1992. A note on the genealogy of research traditions in modern phonology. *Journal of Linguistics*, *28*, 149–163.
Golston, Chris. 1998. Constraint-based metrics. *Natural Language and Linguistic Theory*, *16*, 719–770.
Hale, Mark & Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
Hayes, Bruce, Robert M. Kirchner & Donca Steriade. 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.
Hooper, Joan Bybee. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
Iosad, Pavel. 2012. Representation and variation in substance-free phonology: A case study in Celtic. PhD dissertation, University of Tromsø.
Johnson, Keith, 2007. Decisions and mechanisms in exemplar-based phonology. In Maria-Josep Solé, Patrice Speeter Beddor & Manjari Ohala, eds, *Experimental approaches to phonology*, 25–40. Oxford: Oxford University Press.
Kiparsky, Paul. 1982. Lexical phonology and morphology. In I. S. Yang, ed, *Linguistics in the morning calm*, 3–91. Seoul: Hanshin.
Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook*, *2*, 85–138.
Kiparsky, Paul. 2016. Labov, sound change, and phonological theory. *Journal of Sociolinguistics*, *20*, 464–488.
Liberman, Mark. 1975. The intonational system of English. PhD dissertation, Massachusetts Institute of Technology.
Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry*, *8*, 249–336.
Mohanan, K. P. 1982. The theory of lexical phonology. PhD dissertation, Massachusetts Institute of Technology.
Mohanan, K. P. 1986. *The theory of lexical phonology*. Dordrecht: Reidel.
Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
Palmer, F. R. (ed.). 1970. *Prosodic analysis*. London: Oxford University Press.
Pierrehumbert, Janet. 2000. The phonetic grounding of phonology. *Bulletin de la communication parlée*, *5*, 7–23.
Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Librairie Payot.
Schane, Sanford A. 1973. *Generative phonology*. Englewood Cliffs, NJ: Prentice-Hall.
Selkirk, Elisabeth. 1980. *On prosodic structure and its relation to syntactic structure*. Bloomington: Indiana University Linguistics Club.
Tranel, Bernard. 1981. *Concreteness and generative phonology: Evidence from French*. Berkeley: University of California Press.

Part I

Theoretical frameworks

2
Optimality Theory
Motivations and perspectives

Pavel Iosad

2.1 The basics of OT

The classic version of Optimality Theory, first described in detail by Prince & Smolensky (1993), is a phonological framework that privileges the simultaneous satisfaction of multiple violable constraints by phonological representations over the gradual construction of correct representations from given inputs. At its core, an OT grammar implements a search procedure that finds the surface form that is most compatible with the relevant underlying representation, given the specific properties of the particular language. In OT, these language-specific properties are encoded as a series of rankable violable *constraints*.

2.1.1 Constraints: Con

In formal terms, a *constraint* is most commonly understood as a function that maps from an ⟨input, output⟩ pair to an integer corresponding to the number of violations incurred by that pair. In practice, constraints are most commonly formulated as imperative statements of the form "Assign a violation mark [i.e. increment the return value of the function by 1] for each instance in ⟨input, output⟩ of a structure characterized by some property X". Thus, a constraint against postvocalic stops, which in OT parlance could be written as *V[−son −cont], would penalize any substrings of an output candidate that contain a vowel, followed by a stop. A form like [tata], therefore, incurs one violation of the constraint; a form like [tapata] incurs two; and forms like [taθ] and [saθa] incur zero—in the former case since no stop in the form is postvocalic, and in the latter case since there are no stops in the form at all.[1]

In standard OT, constraints come in two varieties: *markedness* and *faithfulness* constraints. As discussed by Moreton (2004), markedness constraints are distinguished by the fact that the number of violations they assign does not depend on the properties of the input in the ⟨input, output⟩ pair. The constraint *V[−son −cont] is an example: it only refers to properties of the output, and always assigns the same number of violation marks (zero) to a candidate with output [taθa], irrespective of whether the input is [tata] or [taθa]. Markedness constraints, therefore, are statements about the preferred shape of surface representations.

Faithfulness constraints, on the other hand, demand that certain aspects of the input should be preserved in the output. Formally, a faithfulness constraint never assigns a violation mark to the *fully faithful candidate*: a constraint C is a faithfulness constraint if there are no ⟨input, output⟩ pairs such that the input is identical to that output and C assigns a violation mark to the pair. A common type of faithfulness constraint, for example, demands that input and output be identical in the value of some distinctive feature. A constraint like Ident-IO([±continuant]) will therefore assign one violation mark to an input–output pair ⟨/ta**t**a/, [ta**θ**a]⟩ (since the highlighted segment changes its value of [±continuant]) but none to ⟨/tata/, [tata]⟩.

2.1.2 The candidates: Gen

The set of potential output forms produced by the module Gen is, in classic OT, assumed to be both infinite and independent of the properties of the input. The only restriction on Gen admitted in classic theory is the extent of "the basic structural resources of the representational theory" (Prince & Smolensky 1993: 6): in other words, Gen makes available, as potentially corresponding to any input, any phonological object that does not violate the tenets of representational theory.

This independence of outputs from inputs severely restricts the analyst's freedom to account for phonological phenomena in ways other than constraint (re)ranking. If some output is ungrammatical in a language (for a particular input), this cannot be – in classic OT – accounted for by assuming that this derives from some input property that restricts the range of potential outputs. The explanatory burden is thus shifted from restrictions on the set of forms to be considered to the evaluation procedure, described in the next section.

2.1.3 The evaluation procedure: Eval

The fundamental concept of OT is *constraint ranking*. All the constraints are arranged in a relation of *dominance*, which is transitive: in each pair of constraints either one outranks the other or they are unranked with respect to each other, and if A ≫ B and B ≫ C then necessarily A ≫ C. The concept of ranking comes into play in the Eval module, which chooses the correct output from the candidate set offered by the universal mechanism Gen.

In a nutshell, Eval chooses the candidate that has the fewest violations of the highest-ranked constraint. In practice, one constraint is rarely enough to select the "winner", so the procedure is commonly described recursively. For each constraint and a pair of candidates, it is possible to determine if the constraint *favours* one of the candidates: the favoured candidate accrues fewer violations than the disfavoured one. Once the constraints are arranged in ranking order, the candidate set is winnowed by rejecting all candidates that are disfavoured by the highest-ranked constraint, in the sense that there exist other candidates that accrue fewer violations of that constraint. In classic OT, once a candidate is excluded from consideration by some constraint, it can never be a winner (domination is *strict*) – but see Krämer (this volume) for discussion of other alternatives. Those candidates that survive this procedure are passed on to the next constraint in the ranking, and the winnowing is repeated until either the bottom of the ranking is reached or there is only one candidate left.

This interplay of markedness and faithfulness has an important consequence for the scope of the theory. In OT, the concept of a phonological rule – a mechanism that rewrites part of an input string – is largely replaced by that of the unfaithful mapping, whereby the winning candidate is not identical to the input in one respect or another. Unfaithful mappings necessarily violate faithfulness constraints: therefore, they can only be allowed when the relevant faithfulness constraints are outranked by some markedness constraints. This is the basic markedness-over-faithfulness (M ≫ F) schema. Consider the example of *gorgia toscana*: a process in Tuscan varieties of Italian whereby stops become fricatives in intervocalic positions (Giannelli & Savoia 1978, 1980; Giannelli & Cravens 1996; Kirchner 2000).[2] In isolation, input /kasa/ for 'home' is realized as [ˈkaːsa], whereas in a phrasal context after a vowel the initial stop becomes a continuant, often [h]: [la ˈhaːsa] 'the home'. Ignoring for now the assignment of prosodic structure (specifically stress assignment and vowel lengthening) and /s/-voicing, the crucial mapping is /k/ → [h]. This mapping violates a faithfulness constraint that demands the preservation of [±cont] values (Ident[±cont]). In the citation form, this constraint remains unviolated, but in the postvocalic context it cannot prevent the unfaithful mapping. Thus, for instance, Ident[±cont] outranks the constraint *[−son −cont] that militates against surface stop consonants: in other words, it is more important to avoid changes in continuancy (e.g. spirantization) than to avoid the presence of stops in the output. In OT work, this is formalized as in (1).

(1) Faithfulness over markedness

		/kasa/	Ident[±cont]	*[−son −cont]
a.	☞	[kasa]		*
b.		[xasa]	*!	

The diagram in (1), referred to in OT as a *tableau*, demonstrates the derivation of [kasa] from /kasa/. Candidate (a.), which eventually wins, is favoured by Ident over candidate (b.), where the [cont] value of the [k] is changed, acquiring one violation of the faithfulness constraint (marked by the asterisk). Since the fully faithful candidate (a.) is favoured by the Ident constraint, candidate (b.) is knocked out of contention (as the exclamation mark indicates). Although the favouring relationship is reversed for the constraint [−son −cont], that constraint is ranked too low, i.e. below the Ident constraint, to force the choice of candidate (b.)

When the input puts the stop in a postvocalic position, the mapping effected is unfaithful. This must be due to a different markedness constraint, militating against postvocalic stops. This constraint outranks the Ident constraint.

(2) Markedness over faithfulness

				*V[−son −cont]	Ident[±cont]	*[−son −cont]
/kasa/	a.	☞	[kasa]			*
	b.		[xasa]		*!	
/la kasa/	c.		[la kasa]	*!		*
	d.	☞	[la hasa]		*	

In the case of the input /kasa/, neither candidate violates the contextually determined markedness constraint, since neither contains the offending sequence of a vowel and a stop (in these cases the constraint is sometimes said to be *vacuously satisfied*). The evaluation is passed on to the next constraint – Ident – with the same results as in (1). As for /la kasa/, the constraint against postvocalic stops knocks out the faithful candidate and an unfaithful mapping ensues.

It is important to note a non-trivial shift in emphasis compared to rule-based phonology. There is no real analogue to a "faithful mapping" in a rule-based theory: if an output happens to be identical to its input, this is an epiphenomenon of the fact that no rule happens to apply to it, and it does not require a special account. In OT, a faithful mapping requires an explanation just as much as an unfaithful one: it becomes a phonological fact. In fact, the faithfulness of mappings plays an important role in much of the work on learnability (Tesar 2013), which is likely to have been facilitated by the fact that it can be expressed formally within the theory.

However, not all phenomena of interest to the phonologist boil down to unfaithful mappings. A notable class of cases includes structure-building operations such as the construction of prosodic structure. The classic example is syllabification, which is commonly assumed to be driven solely by the interaction of several markedness constraints, such as Onset (penalizing onsetless syllables) and NoCoda (penalizing codas). Faithfulness to syllable structure does not come into the picture: this is catered for by an assumption that faithfulness constraints to syllabic structure are not part of Con. This assumption is not vacuous: it predicts that there can be no lexical contrasts in syllabic affiliation – even if such a contrast were present in the input, the absence of a faithfulness constraint predicts that any underlying specifications would be overridden by the markedness constraints. This prediction is widely assumed to be correct (e.g. McCarthy 2007a), although see Clements (1986); Vaux (2003); and Köhnlein (2016) for potential counterexamples.

The main distinguishing features of the OT framework can be summarized as follows:

- Instead of descriptions of *processes*, OT focuses on descriptions of *desirable outputs*: processes emerge as the result of satisfying these descriptions at the cost of changing inputs to the extent allowed by the particular language;
- Instead of pattern-specific descriptions, violability allows the analyst to formulate the constraints in a more general way: pattern specificity arises from the competing demands of various general constraints;
- The number of "moving parts" in a theory of cross-linguistic variation is reduced: neither the input nor the structure-building mechanism are allowed to be language-specific, with ranking acquiring a crucial role in accounting for variation.

All of these will be treated in more detail in the following sections.

2.2 Some advantages of OT

The "creation myth" of Optimality Theory often sees its flowering as the culmination of a dissatisfaction with the input-oriented, process-heavy framework rooted in the SPE model. In this view, rule-based approaches are vitiated by an emphasis on the properties of the input, which trigger the (non-)application of rules, and formal difficulties with expressing generalizations about the output of rules. It is common to cite a number of developments within rule-based phonology as being important precursors to constraint-based frameworks:

- Morpheme-structure constraints (Stanley 1967; Sommerstein 1974; Shibatani 1973), i.e. statements of what input shapes are allowed in a particular language;
- Conspiracies (Kisseberth 1970), i.e. situations where several formally disparate rules converge on outputs with a particular property (for a worked-out OT example, see Pater 1999);

- Autosegmental phonology (Goldsmith 1976 and much subsequent work), with its emphasis on the description of conditions that a representation must fulfil in order to be licit, and its further developments such as feature geometry (Sagey 1986; McCarthy 1988; Clements & Hume 1995);
- Developments in the analysis of templatic morphology (McCarthy 1979) and prosodic phenomena (Marantz 1982), which privileged descriptions of licit outputs over procedural parsing "directions" (or at least included both components).

The apparent inability of rule-based phonology to deal with these issues in a satisfactory manner is commonly seen as paving the way first for hybrid rule-and-constraint frameworks, starting already with Sommerstein (1974) and later in constraints-and-repairs theories (Paradis 1992; Calabrese 2005), and then for constraint-only formalisms, which include not only OT but also various flavours of Declarative Phonology (Scobbie et al. 1996; Scobbie 1997; Coleman 1998) and Government Phonology.

2.2.1 Factorial typology and harmonic bounding

Several other competitive advantages of OT flow from its computational properties. In classic OT, the universality of Con obviates the need to learn the constraints; the principles of candidate generation also preclude language-specificity in the choice of possible outputs; and adherence to Richness of the Base negates the unavoidably language-specific character of the lexicon. It follows that languages differ *only* in the relative ranking of the constraints, of which there is a finite number. For the analyst, this corollary opens up the tantalizing possibility of doing highly explicit typology. The inputs, potential outputs, and constraints (of which there is a finite number) are all fixed; the Eval procedure is essentially guaranteed to produce an output for a given ranking; it is consequently feasible (at least in principle) to identify the input–output mappings given by all logically possible permutations of constraint orderings.

This enterprise is known as *factorial typology* (since n constraints can be ranked in $n!$ ways). Consider, for instance, our analysis of *gorgia toscana* above. The unfaithful mapping in Tuscan Italian is enabled by the ranking of a markedness constraint (*V[ptk]) over a faithfulness constraint Ident-IO([cont]). This ensures that postvocalic stops change their [±continuant] value to satisfy the markedness constraint. In a toy grammar with two constraints, there are only 2! = 2 permutations: the reverse ranking produces a faithful mapping with the stops intact.

(3) No spirantization with reverse ranking

		/gato/	Ident-IO([cont])	*V[ptk]
a.	☞	[gato]		*
b.		[gaθo]	*!	

It is thus predicted – correctly if rather trivially – that language with and without postvocalic spirantization should both exist: Standard Italian is an example of the latter.

Now consider the fact that in Tuscan Italian geminate stops resist spirantization: contrast [kaˈɸiθo] 'understood' with [ˈskritːo] 'written'. This is a typologically common phenomenon that has received a variety of explanations (e.g. Schein & Steriade 1986; Kirchner 2000; Honeybone 2005). For the purposes of the argument, we can assume a rather descriptively formulated IDENT-IO$_{gem}$ constraint family, which assigns a violation mark to all surface geminates that undergo featural change, and introduce it into our analysis of Tuscan.

In (3), this constraint is vacuously satisfied by both candidates of interest, since there are no geminates in the input. Consider now an input with a geminate stop:

(4) No spirantization with geminates

		/skrit:o/	IDENT-IO$_{gem}$([cont])	*V[ptk]	IDENT-IO([cont])
a.	☞	[skrit:o]		*	
b.		[skriθ:o]	*!		*

The grammar in (4) is a superset of that in (2), but it also provides for geminate inalterability. It also includes three constraints rather than two, so technically there are 3! = 6 potential rankings. However, the number of possible input–output *mappings* is not equal to the number of *rankings*. Consider a grammar of Standard Italian with the three constraints:

(5) No spirantization in any context

				IDENT-IO$_{gem}$([cont])	IDENT-IO([cont])	*V[ptk]
/gato/	a.	☞	[gato]			*
	b.		[gaθo]		*!	
/skrit:o/	a.	☞	[skrit:o]			*
	b.		[skriθ:o]	*!	*	

Note that the ranking in (5) is part of the six-member factorial typology for the three constraints, although there is no actual evidence for the relative ranking of the two versions of IDENT-IO([cont]): the two constraints are never in conflict. In fact, it can easily be verified that *reversing* their ranking has no effect on the outcomes in (5). Similarly, in a grammar where *V[ptk] dominates both of the faithfulness constraints, the relative ranking of the latter two is unimportant, because the high rank of markedness imposes an unfaithful mapping. In fact, the six logically possible rankings produce only three types of mappings, as summarized in Table 2.1.[3]

In particular, note that when the more general faithfulness constraint dominates markedness, it also ensures a faithful mapping in cases covered by the more specific constraint, so that even if the latter is strictly speaking dominated by markedness, no unfaithful mapping can occur. The explicit prediction that follows from this (toy) factorial typology exercise can be reformulated as an implicational universal: spirantization in geminates implies spirantization in singletons but not vice versa. Another way to look at the predictions is in terms of

Table 2.1 Factorial typology of spirantization

Ranking	*Spirantization in singletons*	*Spirantization in geminates*
M ≫ F, PF	yes	yes
PF ≫ M ≫ F	yes	no
PF, F ≫ M	no	no
F ≫ M, PF	no	no

possible mappings. The *possible* input–output mappings can be visualized in terms of a cline (similar to the traditional lenition trajectories):

- Spirantization everywhere: /| tt t → θ/;
- Spirantization in singletons only: /tt | t → θ/;
- No spirantization: /tt t | θ/.

The impossible mappings are as follows:

- The chain shift: /tt → t → θ/;
- The "saltation" (Lass 1997; Hayes & White 2015): /tt → θ/ with /t/ unchanged.

These impossible mappings are said to be *harmonically bounded*, because there is no ranking under which they are better ("more harmonic") than some other candidate, which thus set an "upper bound" to the losing candidate's harmony under the available set of constraints.

Note that this type of interaction between specific and general constraints emerges with no additional stipulation from the universally defined mechanism of constraint interaction. More specifically, it emerges when two (or more) constraints stand in a *stringent* relationship, meaning that the set of violations assigned by one constraint is always a superset of the set of violations assigned by the other: in our case, any candidate that violates Ident-IO$_{gem}$ must by necessity violate the simpler Ident-IO (at least) as many times. Constraints in such a relationship do not, strictly speaking, conflict with each other, so their mutual ranking is not important: their ranking vis-à-vis other constraints, however, matters for the outcome.

This example shows, however briefly, the basic mechanism whereby factorial typology enables the analyst to do phonological typology. Typology is primarily seen in terms of possible input–output mappings.[4] Factorial typology, at least in principle, makes explicit – and therefore falsifiable – predictions about possible phonological grammars. It is therefore no surprise that OT has found particularly widespread use in domains for which the typological parameters are relatively well understood, such as syllable structure (already in Prince & Smolensky 1993), weight (e.g. Morén 2001), and metrical typology (e.g. McCarthy & Prince 1995a; Hyde 2001; Alber 2005). Applications of these methods in other areas (notably segmental phonology) have been less prominent, although cf. Causley (1999) and de Lacy (2006) for some results.

Of course, factorial typology is promising as a potential account of the full range of cross-linguistic variation only if the set of constraints Con is finite and at least potentially fully discoverable. In classic OT, this is achieved via the assumption that Con is not just finite but also *universal*. If this is correct, then all constraints are present in the grammars of all languages, and their (apparent) (in)activity should be accounted for by reference to their ranking.

The assumption of universality is not logically unavoidable: it is perfectly conceivable that constraints might be constructed by the learner as part of the acquisition process (cf. for instance Hayes & Steriade 2004; Pulleyblank 2006; Archangeli & Pulleyblank, this volume). An OT framework with such emergent constraints is, of course, viable, but it loses the possibility of accounting for cross-linguistic differences *solely* by reference to constraint ranking via factorial typology. Even so, it retains some other properties that made the framework attractive in the first place.

2.2.2 Increased generality

Compared to other constraint-based frameworks, such as Declarative Phonology, OT's embrace of constraint violability offers the promise of increased generality of explanation. If constraints are inviolable, it is likely that every specific pattern requires a precisely described constraint which refers exactly to the contexts where the relevant phenomenon is in evidence. With OT's violable constraints, multiple patterns can be described with a rather smaller number of more general constraints. A candidate may violate some of these constraints and yet still be the winner, by virtue of constraint ranking or vacuous satisfaction.

One class of cases that is common to cite as exemplifying this advantage is referred to as *emergence of the unmarked* (McCarthy & Prince 1994). In this situation, evidence for some ranking of constraints is available only when some more highly ranked third constraint is, for some reason, inactive. One situation where this arises is when the third, dominating constraint is a faithfulness constraint, so that the interaction of the constraints it outranks only becomes evident when faithfulness is vacuously satisfied.

By way of example, consider a language whose surface inventory possesses both [i] and [u]. The existence of these vowels is militated against by constraints on the co-occurrence of their respective features, call them for brevity *[−bk +hi] and *[+bk +hi]. The fact that the vowels contrast in the language indicates that relevant faithfulness constraints (such as, say, Ident-IO[±bk]) dominate the two markedness constraints. For this reason, the relative ranking of the two markedness constraints cannot be easily recovered, since they do not participate in choosing a winner candidate. Now imagine, however, that the language also possesses a process of vowel epenthesis, which by definition involves the appearance of a vowel that does not correspond to anything in the input. In this case, Ident-IO is vacuously satisfied, and the choice of the quality of that vowel falls to other constraints – perhaps to the two markedness constraints just identified. Now if the epenthetic vowel happens to be [i], this provides evidence that the constraint against surface [u] outranks the constraint against surface [i]: the relative (un)markedness of the two vowels is "submerged" by the high ranking of faithfulness, but emerges if the latter is rendered inactive. Crucially for the question of generality, note that there is no need to introduce any mechanisms specific to the situation of epenthesis: all the work is done by the ranking, using constraints that, under OT assumptions, are independently needed to describe the language's surface contrasts. This increased generality of constraints is part of the reason that OT is at all viable as an instrument of cross-linguistic explanation, making typological arguments of the sort exemplified in section 2.2.1 possible.

2.2.3 OT and markedness

The OT mechanism provides an explicit formalization of the age-old idea – going back at least to Baudoin de Courtenay – that the form of speech utterances represents a compromise between the needs of the speaker (such as minimizing effort) and the needs of the hearer (such as ambiguity avoidance). In OT, the "needs of the speaker" are largely expressed via markedness constraints, which tend to require that surface representations have certain properties (such as having only open syllables or lacking front rounded vowels) and, by implication, that they do not have some other properties (such as syllable codas and front rounded vowels). In itself, such a requirement has little to do with "markedness" as understood in pre-OT literature with reference to properties such as typological distributions or particular types of phonological behaviour. Nevertheless, as the corpus of OT analyses grew, it became apparent that the (possibly universal) set of markedness constraints available to learners must include constraints with a clear affinity to the phonetic factors commonly implicated in accounting for (un)markedness in phonological behaviour.

The establishment of the link between "phonetic knowledge" (Kingston & Diehl 1994) and phonological grammar opened up a significant field of inquiry. Accounts of markedness effects proved notoriously difficult to incorporate directly into rule-based accounts. Markedness-based asymmetries could be introduced via additional submodules of the grammar (such as the "markedness conventions" of Chomsky & Halle 1968 or the "grounding" conditions of Archangeli & Pulleyblank 1994) or through language-specific representational underspecification (Steriade 1987; Archangeli 1988), which eluded comprehensive cross-linguistic theorizing. In OT, the theory of markedness is part and parcel of the theory of constraints; and a theory of Con is perhaps the most important part of the theory of grammar, since all other components of the mechanism (the Gen mechanism, the rich base, the evaluation mechanism) are essentially fixed.

The OT formalism itself does not put a restriction on the substantive content of markedness constraints. The default position is to make them refer to orthodox pieces of phonological structure, such as features and suprasegmental constituents; however, various authors have proposed that they could also refer directly to phonetic properties such as formants (Flemming 2002) or cues (Steriade 2001), articulatory effort values (Kirchner 1998), and properties of entire (sub)inventories (Padgett 2003). In all cases, however, OT offers clear advantages to any theorist who wishes to account for markedness effects via some property of Universal Grammar – although this aspiration is not universally shared (e.g. Hale & Reiss 2008).

2.2.4 Conspiracies and the "duplication problem"

Another commonly cited advantage of OT is the resolution of the "duplication problem", whereby apparent restrictions on the form of underlying representations are reproduced in dynamic alternations. Consider the case of velar fronting in Modern Standard Russian. Generally, sequences of a non-palatalized velar [k ɡ x] and a front vowel are disallowed in underlying forms of morphemes.[5] In parallel, a rule of velar palatalization ("velar fronting" in the tradition initiated by Halle 1959) maps the sequences /ki ɡi xi ke ɡe xe/, when they arise via (some kinds of) suffixation, to /kʲi ɡʲi xʲi kʲe ɡʲe xʲe/: /potolok/ 'ceiling' but [patalˈkʲi] 'ceiling-NOM.PL', [patalˈkʲe] 'ceiling-LOC.SG'.

In a serial framework, the first restriction is either left unexpressed – treated as an accident of history – or enforced by some mechanism specific to underlying representations, whereas the second appears to necessitate a phonological rule, although the outcome of these two mechanisms is identical. In mainstream OT, neither of these techniques is allowed because of the postulate of Richness of the Base, which bans positing any restrictions on inputs, whether accidental or in a dedicated submodule of the grammar. Under this régime, the apparent restriction on underlying forms is illusory: it is incumbent on the grammar to rule out a situation whereby (potential) underived forms containing the offending structure are mapped to the surface faithfully. This can be done by a proper ranking of some markedness constraint militating against this structure – which is the same mechanism needed to enforce alternations. Thus, OT proponents argue that restrictions on underived forms and alternations are accounted for by a single mechanism.

Another area in which OT is claimed to excel is resolving the issue of "homogeneity of target, heterogeneity of process", sometimes known as the issue of "conspiracies" (Kisseberth

1970; Kenstowicz & Kisseberth 1979). The classic case here is Yowlumne (Yawelmani Yokuts), in which several phenomena – rules of vowel shortening and epenthesis, as well as additional conditions on syncope and apocope rules – all "conspire" to produce a set of surface forms with no complex codas. In an input-oriented framework, expressing this requires reference to the output of the rule, which is not possible without additions to the formalism. In OT, on the other hand, the homogeneity of target is directly expressed by the relevant constraint, whereas heterogeneity of process arises as a consequence of differences in the shape of inputs (which render some faithfulness constraint vacuous) and the ranking of faithfulness constraints.

2.2.5 Computational advantages

The area of phonological research that has probably seen the most significant advances compared to the pre-OT results is the study of learning algorithms for phonology. With the advent of OT, problems such as the learning of phonotactic restrictions, alternations, and underlying representations have been subject to mathematically explicit analyses, clarifying both the limits of the OT formalism and the possible ways in which a phonological system may be acquired by a learner.

It became clear rather early on that the use of OT formalism as such does not, in principle, afford significant computational advantages, as the problem of generating the set of winners was shown to be NP-hard (e.g. Eisner 1997; Wareham 1998) – that is to say, there is no difference in computational complexity between OT and standard implementations of rule-based phonology (see Heinz 2011b for an overview of these issues).

Nevertheless, the use of well-understood optimization techniques makes it possible to offer learnable versions of OT grammars. This is largely thanks to the fact that harmonic bounding ensures only that the notionally infinite candidate set contains a smaller subset of "viable" candidates, as there is no need for the grammar to ever consider a large subset of candidates that can never be winners (Seeker & Quernheim 2009; Riggle 2009; Heinz et al. 2009). Issues related to computational complexity, tractability, and suitable algorithmizations continue to be debated in the literature: for instance, they play an important role in some scholars' endorsement of Harmonic Grammar, with its constraint weighting, over OT's logic of strict domination (e.g. Pater 2009b; Potts et al. 2010) – although see Magri (2013) for a rejoinder, and Krämer (this volume) for more discussion.

One area where the advent of OT has undoubtedly led to significant progress is the study of learnability. Algorithms with well-understood properties have been developed to make progress with phonotactic learning (the acquisition of the ability to distinguish felicitous and infelicitous surface forms), the resolution of structural ambiguity (choosing the correct candidate among a set of candidates that do not differ visibly but have different structural parses), resolving the subset problem (choosing the most parsimonious grammar out of the set of grammars consistent with the surface data), and – to a somewhat lesser extent – the learning of underlying representations and of morphophonological alternations.

One basic idea in OT learnability work is the notion of *constraint demotion* (Tesar & Smolensky 2000, Chap. 3). A constraint is demoted if the learner encounters a datum that cannot be accommodated within the grammar they have arrived at (an *inconsistency*). Specifically, it is demoted below a constraint that allows the "correct" candidate to emerge as the winner (but not further down the ranking). This *inconsistency resolution* can be leveraged in a variety of ways, not just for phonotactic learning, but also, for instance, for the acquisition of lexically specific phonological phenomena (Pater 2009a).

Constraint demotion is able to detect inconsistencies, but it is not robust to errors in the data, since any inconsistent datum triggers a reranking. An alternative approach is the Gradual Learning Algorithm (Boersma & Hayes 2001), which uses ambient data not to effect a full-on change of ranking but rather to change the probability of a particular constraint being ranked in a particular way, thus ensuring that the influence of inconsistent data is proportionate to their frequency (in particular, the prediction is that since true errors are rare, they will not unduly influence the acquisition process). The downside of this robustness, however, is that the Gradual Learning Algorithm is unable to detect global inconsistencies (i.e. inconsistencies that arise from more than a pairwise ranking of constraints), which are useful in resolving structural ambiguity (Tesar 2004).

All these, and other related issues, continue to be the subject of active research. The construction of explicit algorithms with well-understood computational properties promises to close the gap between phonological theory and many broader concerns in cognitive science, as well as clarify the scope of phenomena that phonological theory should – or indeed can – be concerned about. The advent of OT has played a significant role in the development of this highly necessary work.

2.2.6 Quantitative gradience

Another area of active research in OT concerns the encoding of quantifiable, gradient phonological phenomena. This is facilitated by several properties of the OT architecture.

First, factorial typologies make explicit predictions as to what phenomena are possible (including, crucially, how different phenomena may interact). Second, since the number of grammars generated by reranking a finite number of constraints is also finite, the number of grammars including a particular output (or constellation of outputs) can be estimated. Thus, for instance, Multiple Grammars Theory (Anttila 1997) uses the insights of factorial typology to account not just for the contextual restrictions on variable processes but also for the quantitative aspects of variation: it predicts that the frequency of a particular variant is proportional to the share of grammars allowing that variant in the factorial typology.

More sophisticated results can be achieved by a more nuanced approach to constraint ranking. Anttila (1997) shows that non-trivial predictions about variation can be made if a "grammar" is allowed to include partial rather than total ordering of constraints (see Krämer, this volume, for more discussion). Another example is Stochastic OT, where the rankings of constraints for each utterance are determined probabilistically, deriving non-trivial quantitative predictions for within-speaker variation.

Gradient phonological knowledge is also apparent in continuous phonotactic effects. Phonotactic knowledge is apparent in many phenomena, ranging from over- or under-representation of certain patterns relative to chance level, effects of well-formedness in production and perception experiments, loanword adaptation, and so on. Many aspects of this knowledge appear to be gradient, in that it is possible to distinguish between "degrees" of well- or ill-formedness (e.g. Schütze 1996). Many proponents of OT take seriously the proposition that this gradience is not "just" a performance effect, but instead should be derived from the same mechanisms as those that underlie categorical grammatical phenomena (e.g. Hayes 2000; Coetzee 2008). Once again, OT provides the means of quantifying such patterns, either through an inspection of the numerical consequences of categorical constraint rankings or introducing a stochastic element into the evaluation procedure.

2.3 Challenges and issues

It will be noted that many of the arguments advanced for OT are ones of theoretical elegance rather than empirical coverage. This is, in principle, not surprising, given that it may be difficult to distinguish the *empirical* coverage of OT from that of other theories, since most of them can be shown to describe (sub)regular relations (Heinz 2011a). In fact, the explicitness of predictions made by OT has uncovered a number of serious challenges to its status as a fully adequate theory of phonological competence. Other issues have arisen as a consequence of choices made by analysts within and outwith the OT tradition in terms of focus.

2.3.1 Opacity

Opaque interactions have been perhaps the most prominent empirical problem for OT. In the classic typology opaque processes can be described in terms of either overapplication (application in the absence of the context on the surface) or underapplication (non-application despite the presence of the triggering context). Both of these are problematic for classic OT, but especially overapplication, because in that case the desired winning candidates are harmonically bounded and thus predicted to *never* win.

A classic case of overapplication is the interaction of palatalization and syncope in Bedouin Arabic, where palatalization of velars can be counterbled by syncope targeting medial open syllables:

(6) /ħaːkim-iːn/ 'rulings' → /ħaːkʲimiːn/ → [ħaːkʲmiːn]

Assuming for the sake of the argument that palatalization violates a constraint Ident-IO[±bk] and is triggered by a constraint Agree[±bk], requiring that a consonant and a following vowel have the same [±bk] value, whereas syncope is triggered by a constraint ranking referred to as FtStruc (expressing a preference for disyllabic feet), we can attempt to construct the following tableau:

(7) Harmonic bounding in Bedouin Arabic counterbleeding opacity

	/ħaːkimiːn/	FtStruc	Agree[±bk]	IdentIO[±bk]	Max
a.	[ħaːkimiːn]	*!	*		
b. ☞	[ħaːkmiːn]				*
c.	[ħaːkʲimiːn]	*!		*	
d.	[ħaːkʲmiːn]			*!	*

The intended winner, candidate (d.), is harmonically bounded by (b.): they both undergo the unfaithful mapping involving deletion, but the candidate with both palatalization and syncope cannot win, since it incurs entirely gratuitous violations of the IdentIO constraint. Thus, parallel OT makes the prediction that counterfeeding opacity should be impossible. This is more than a little problematic, since the existence of such mappings is probably the most significant result of generative phonology, setting it apart from most if not all other phonological frameworks.[6]

Responses to the opacity problem in the literature have been varied. Frameworks such as Sympathy Theory (McCarthy 1999) and Virtual Phonology (Bye 2001) enrich the representational arsenal by reintroducing the possibility of reference to non-surface forms. A more constraint-focused approach is offered by Comparative Markedness (McCarthy 2003), which allows constraints to distinguish the status of various candidates. Several approaches exploit the frequent relationship between opacity and morphology. One example here is Cophonology Theory (Orgun 1999; Orgun & Inkelas 2002; Inkelas 1998), which allows all affixation to trigger morphologically specific phonological effects. Another option is allowing morphologically related words to influence the phonological form of each other, as in Output–Output Correspondence (Benua 1997) and Optimal Paradigms (McCarthy 2004a). Stratal approaches (Kiparsky 2000; Bermúdez-Otero 2011, this volume) use the insights of Lexical Phonology and Morphology (Kiparsky 1982, 1985) to provide a restricted theory of morphological influence on phonological processes. Finally, theories such as OT-CC (Optimality Theory with Candidate Chains) and Harmonic Serialism abandon the principle of unrestricted Gen in favour of a stepwise derivation that restricts the "distance" an output may diverge from its input (McCarthy 2007a, 2008b; McCarthy & Pater 2016).

Another characteristic response of OT proponents to the issue of opacity is essentially a denial of its reification as a single phenomenon that is problematic for OT. For instance, Baković (2007) proposes a revised typology of opaque generalizations and argues that OT is more suited to dealing with certain classes of opaque phenomena than rule-based theories, while Łubowicz (2012) proposes a parallel OT account of a large class of underapplication phenomena without purporting to solve the entire "opacity" problem.

2.3.2 Representations

The advent of OT has coincided with a retreat from much work in representational theory characteristic of the late 1980s. This is particularly true for the study of subsegmental representations, such as feature geometry and underspecification. To a certain extent this may perhaps be viewed as a matter of a new swing of Anderson's (1985) "representation/computation pendulum", whereby OT's focus on the computational device of constraint ranking as an explanatory factor in phonology has led to a fall-off in focused representational work. Architecturally, the OT algorithm does not impose strict logical requirements on the representational properties of inputs and outputs. In practice, however, the emphasis on constraint ranking as the sole explanatory mechanism encourages the use of commonly agreed, cross-linguistically invariant featural representations – which, in practice, has tended to mean the SPE feature set (Chomsky & Halle 1968).

The reasons for the decreased emphasis on underspecification are different. With the premium that OT puts on minimizing the number of violations, a common assumption in much of the literature is that the learning mechanism is geared to produce a set of inputs that can be fed into the correct ranking to produce the right outputs with as few violations as possible. Since the change from an underspecified input to a specified output most often involves the violation of a faithfulness constraint,[7] even when there is no alternation, a fully specified input accrues fewer violations than an underspecified one, in a process referred to by Prince & Smolensky (1993) as "Stampean occultation" (see, however, Krämer 2012; Tesar 2013 for critical discussion of this idea). This is shown in (8). For an output form [da] that does not show any alternations, assuming an identical input /da/ means only markedness constraints are violated in the input–output mapping. On the other hand, assuming an input /Da/, with the first segment underspecified for [±voi], means that in addition to those markedness violations the mapping will also incur a violation of faithfulness, because of the insertion of a [±voi] specification absent in the input. Therefore, in the absence of alternation evidence full specification is preferred.

(8) Stampean occultation: output non-alternating [da]

	Dep[voi]	*[+voi −son]
a. ☞ /da/ ~ [da]		*
b. /Da/ ~ [da]	*!	*

Under this régime, input underspecification can only be countenanced if the option of full specification is not available for some reason. It has been proposed (Inkelas 1994; Krämer 2000) that this is necessary in the case of ternary contrasts, where full specification cannot be shown to derive the correct behaviour. For instance, in Île de Groix Breton (Ternes 1970; Krämer 2000) initial obstruents in lexical items demonstrate three kinds of behaviour:

- Voiceless in isolation, triggers of regressive assimilation in sandhi;
- Voiced in isolation, triggers of regressive assimilation in sandhi: [**b**aːk] 'boat', [atʃypa**ʒ** **b**aːk] 'boat crew';
- Voiced in isolation, undergoers of bidirectional devoicing in sandhi: [**b**ənak] 'any', [atʃypa**ʃ** **p**ənak] 'any crew'.

The behaviour of the first two classes can be derived if they are underlyingly specified as [−voi] and [+voi] respectively, which means the third set cannot have either specification; instead, Krämer (2000) argues, it is underlyingly underspecified for [±voi] and receives the voicing specification via an interaction of markedness constraints regulating the distribution of [±voi] on the surface. Here, underspecification is used as a device to derive ternary behavioural distinctions, when an underlying binary distinction is not analytically viable; this is quite different from the use of underspecification as a representational device to express (lack of) contrast current in much pre-OT work.

It would not be fair to say that work in OT has been entirely unconcerned with questions of phonological representation.[8] In fact, representational questions are crucial to the operation of numerous types of OT analysis. For instance, an important early debate concerned the question of containment vs. correspondence. In much early OT work, the relationship between input and output was a matter of simple containment. This precluded operations such as deletion or insertion: elements appearing to have undergone deletion were assumed to have remained unparsed prosodically (violating constraints of the family Parse), whereas epenthetic material was assumed to represent empty structural positions unfilled by other material (violating constraints of the family Fill). In Correspondence Theory, on the other hand, inputs and outputs were represented separately and their elements related via a many-to-many correspondence relationship: this allowed operations such as deletion (input element with no output correspondent), insertion (output element with no input correspondent), coalescence (multiple input elements corresponding to a single output), and fission (multiple output elements corresponding to a single input). The correspondence relationship can also be extended to pairs of representations other than input–output, including base-reduplicant (McCarthy & Prince 1994), segments in morphologically related forms ("output–output correspondence"; Benua 1997), and surface segments in the same form, as in Agreement by Correspondence (Gunnar Ólafur Hansson 2001; Rose & Walker 2004). The question is clearly a representational one, albeit driven largely by OT-specific concerns rather than the pre-OT preoccupation with contrast. (The question is, incidentally, unresolved: while Correspondence is often the silently assumed option, a line of recent work has resurrected

Containment in the guise of Coloured Containment; van Oostendorp 2007; Trommer 2011; Trommer & Zimmermann 2014).

Representational work in OT can also be motivated by the fact that traditional representations may create empirical problems within the OT computational system. An example is the theory of Feature Classes (Padgett 2002), devised to solve a "pathology" – an undesirable prediction of the factorial typology – that arises with representations based on autosegmental mechanisms such as spreading, namely the so-called "sour grapes" problem: with traditional representations, the factorial typology includes grammars where multiple features spread except when something prevents one of them from spreading and consequently *all* the features fail to do so. It is also possible to find examples where representations are leveraged to achieve certain constraint violation profiles, which are then in turn utilized to build particular factorial typologies: these can be either traditional autosegmental representations (Causley 1999; Iosad 2012) or "bespoke" ones, as in the "*xo* Theory" of de Lacy (2006).

Many such representational devices proposed by OT are driven by the non-serial nature of the computation. For instance, the inherently iterative character of autosegmental spreading sits rather poorly with the fell-swoop OT grammar, and thus a variety of options were offered to capture the same insights using representational rather than derivational options, such as Optimal Domains theory for tone (Cassimjee & Kisseberth 1998), strictly local spreading in vowel harmony (Ní Chiosáin & Padgett 2001), and Binary Domains Theory (Jurgec 2010). It remains to be seen to what extent such representational reimaginings of serial derivation will remain relevant with the reintroduction of serial derivation in more recent versions of the framework (McCarthy 2010).

Finally, it must be emphasized that OT is not inherently inimical to more traditional representational work, as shown, for instance, by the existence of OT analyses making use of featural underspecification motivated by contrast (Hall 2007; Mackenzie 2013; Youssef 2015). Similarly, Hyde (2009) provides a careful comparison of a parallel and a serial account of metrical stress patterns and identifies a well-known representational device (Weak Layering, i.e. the possibility of unfooted syllables within a prosodic word) as the source of some predictions rather than the OT mechanism in use.

2.3.3 Overgeneration and explanatory power

Another area in which the explicitness of OT's predictions came to be presented as a challenge for the theory is the question of overgeneration. The OT formalism is powerful enough to accommodate as real phonological processes input–output mappings that are highly implausible. Recent work in computational complexity shows that phonological patterns are at most regular mappings, and even likely to be restricted to a subregular class (Heinz & Idsardi 2011; Heinz 2011b), which means that most of the computational power of OT (which may extend to context-free grammars – Frank & Satta 1998; Karttunen 1998) goes unused. This creates a serious overgeneration problem, given the default OT position that constraint ranking and harmonic bounding is the only mechanism available to exclude certain mappings (apart from the vaguely defined restrictions on Gen).

Even these considerations aside, it quickly became apparent that "homogeneity of target, heterogeneity of process" – initially seen as an advantage – could be problematic. The classic case is final devoicing. In a rule-based theory, its existence is accounted for by the possibility of a rule that does just that – maps an input voiced obstruent to a voiceless one. The fact that certain other processes in the same position appear to be unattested can, if desired, be accounted for by assuming that the relevant rules are impossible: for instance, Kiparsky (2008) argues that a rule of final obstruent voicing is an impossible one.

In OT, however, unfaithful mappings occur when a markedness constraint dominates some faithfulness constraint. The markedness constraint (or set of ranked constraints) identifies the structure that is to be avoided – such as a word-final voiced obstruent – but the nature of the chosen repair depends on what faithfulness constraint is low ranked: this is the essence of the separation between constraints and repairs. This has long been known to lead to overgeneration: for instance, a word-final voiced obstruent could be repaired by any number of processes, including devoicing, nasalization, outright deletion of the offending segment, epenthesis of material that takes the offending segment out of word-final position, etc., many of which appear to be poorly attested (Steriade 2001; Lombardi 2001).[9] This is known as the too-many-solutions problem, and numerous solutions have been proposed to deal with it, from careful construction of the constraint set Con to match the typology (e.g. Lombardi 2001) to the introduction of new constraint types – e.g. "targeted constraints" (Wilson 2001) or constraints on input–output mappings (Blumenfeld 2006) – to the embrace of gradual derivations (e.g. McCarthy 2007a), which excludes a class of extraneous solutions that arise due to the fell-swoop nature of derivation in classic OT.

Where opacity presents an undergeneration problem, and thus an important empirical challenge to OT, overgeneration is an issue that endangers the theory's claim to theoretical elegance. As discussed in section 2.2, a large part of OT's initial attraction was precisely the greater explanatory power it appeared to offer over the highly powerful rule-based derivations. The explicitness of OT factorial typologies has demonstrated the potential for numerous "pathologies", i.e. predictions that do not correspond well with the set of attested patterns, such as the "midpoint pathology" (predicting that constraints may conspire to push stress as close as possible to the middle of a form), or the "sour grapes" problem mentioned above. It is commonly accepted that, *ceteris paribus*, an analysis that makes fewer (or no) pathological predictions is to be preferred – although it might be objected that the parity of the *cetera* necessary to effect this comparison can be difficult to achieve.

2.3.4 Modularity

The parallel computation of classic OT sits rather awkwardly with the feed-forward cognitive architecture assumed in most generative work. In a feed-forward architecture, the computational mechanisms involved in the production and perception of a phonological expression may operate on that expression in several passes. It is usual to assume at least one pass of some kind of strictly phonological component – or several, in interactionist frameworks such as Lexical Phonology and Morphology. In addition, there may be a gradient component sensitive to factors such as frequency, or perhaps a "usage" component responsible for socially determined variation.

In parallel OT, only a single pass of the computation is available. This may create empirical problems, as in the case of opacity, but it may also be seen as a conceptual difficulty. As emphasized by Scheer (2010, 2011), the single pass gives rise to the "scrambling trope", where all the factors that influence a particular phonological expression have to interact with one another within a single ranking. Examples of this include direct reference to articulatory or acoustic measures, as in much "phonetically based phonology" work, lexical frequency, morphological affiliation (e.g. whether a particular segment belongs to an affix or a root), lexical stratification (e.g. the status of a morpheme as borrowed or native), and so on. This, of course, massively expands the set of interactions that may be predicted to be possible, opening the way, for instance, to morphologically conditioned phonetics, which many proponents of generative theories of phonology would consider to be impossible (Bermúdez-Otero 2010).

The "scrambling trope" is *not* a logically necessary component of OT, as demonstrated by the existence of proposals that explicitly restrict what kind of information can interact; see, for instance, van Oostendorp (2007) and Bermúdez-Otero (2012) for proposals explicitly couched in a modular framework; as Bermúdez-Otero (2010) points out, some frameworks restrict this more implicitly, via excluding some types of constraints from Con, giving Bidirectional Phonology (Boersma & Hamann 2008; Hamann 2009) as an example. Nevertheless, many widely accepted proposals do rely on this kind of "mixing of levels" to resolve important issues in OT, and the possibility (and desirability) of a fully modular OT also remains a live issue at this point.

2.4 OT as a theory of symbolic computation

To sum up this discussion, it is worth revisiting the issue of what it means for a phonological theory to possess explanatory adequacy. Historically, the appeal of OT has only partially been based on better empirical coverage compared to earlier frameworks: to the extent such empirical advances were made – we may mention learnability and the analysis of quantitative aspects of phonological knowledge – they appeared some time after the broader adoption of the theory. In bread-and-butter areas of phonological analysis, OT's advantages were largely perceived as conceptual rather than empirical, and they came with trade-offs in the shape of empirical challenges, perhaps most notably opacity (cf. Vaux 2008).

The future development of OT will be determined by a number of competing pressures. First, like much of formal phonological theory, mainstream OT faces the challenge of justifying its scope and the quality of the underlying data in the face of empirical advances made in the laboratory (Cohn, Fougeron & Huffman, this volume) and in quantitative studies (Hammond, this volume). For OT, this is both a challenge – as the empirical foundations of much of phonological theory become increasingly problematized – and an opportunity, given its pedigree and ethos of incorporating quantitative analysis into phonological grammar. Second, the empirical problems – such as opacity and some important pathologies – have not yet achieved a commonly accepted resolution. These problems continue driving the theory forward, motivating developments such as the (re)introduction of serial derivation in various guises, a shift to constraint weighting, and further work on constraint architecture (Ramsammy, this volume; Krämer, this volume).

Perhaps the most important question that still requires an answer is the scope of the OT computation. At heart, OT is not a proprietary theory of phonology, but a rather general decision-making algorithm. It thus appears to be suitable for the analysis of a broad range of phenomena within a single mechanism. It is this possibility of a single solution for a whole host of issues that appears to have played such an important role in its adoption. Yet many of the issues OT purports to resolve do lend themselves to other remedies. For instance, as discussed in section 2.2.4, OT offered a solution to the so-called "duplication problem"; yet this can only be counted in its favour if one accepts that the "problem" exists and is relevant for an account of phonological knowledge – see Paster (2012) for an argument that it is not, but is instead better understood in a diachronic context. Similarly, many phenomena presented as insurmountable empirical problems that require the introduction of some theoretical device or another can be reanalyzed with a change of assumptions. For instance, in discussing the case of counterbleeding opacity in Bedouin Arabic palatalization, McCarthy (2007a) dismisses an account based on "coalescence" (i.e. the preservation of the [−back] feature of the palatalization trigger through its realization on the consonant), but only adduces conceptual rather than empirical arguments against such an analysis. Another example is the treatment of frequency-sensitive exceptionality, as in the paradigm case of English *comp*[ə]*nsation* vs. *cond*[ɛ]*nsation* (Chomsky & Halle 1968) – treated by Pater (2000) as requiring lexical indexation, this phenomenon has been reanalyzed in an OT framework, yet without recourse to indexation, by Bermúdez-Otero (2012) through a reconsideration of lexical insertion processes.

In sum, I suggest there are two viable directions for the development of OT, which we might call "expansionist" and "minimalist". Under the "expansionist" view, OT and its relatives such as Harmonic Grammar are promising because they offer the possibility of a grand theory of all aspects of phonological knowledge, including not just traditional areas of concern to phonologists such as phonotactics and morphophonological alternations (including typological aspects – see, for instance, Pater 2016 for discussion of the typological merits of constraint weighting) but also quantitative aspects of phonological behaviour, external interfaces, exceptionality, and so on. Under the "minimalist" view, on the other hand, OT occupies a rather more restricted, but perhaps better-defined, place in a theory of phonology, alongside well-articulated theories of interactions between phonology and phonetics, phonology and morphology, phonological computation and lexical access, perhaps also phonological representations or the interaction of phonological and social knowledge, and so on (for some examples, see Blaho 2008; Bermúdez-Otero 2012). It remains to be seen which, if any, of these directions prevails in work on OT; both of them, however, crucially depend on further development of the kind of empirical and theoretical research informed by the questions that OT has raised that is described in the following two chapters of this volume.

2.5 Further reading

The original paper in which OT was introduced (Prince & Smolensky 1993; also published in 2004 by Blackwell) is fairly accessible and provides a good introduction, although many of the specific technical devices it introduces (such as fixed constraint rankings or containment) have since been effectively abandoned or at least problematized. Many key notions, notably "emergence of the unmarked", correspondence, and alignment, are discussed in the influential papers by McCarthy & Prince (1995a, 1995b, 1999). A key paper in the treatment of underspecification and the shape of inputs is Itô, Mester & Padgett (1995).

An article-length introduction to OT is provided by McCarthy (2007b), while McCarthy (2002) offers a survey of the state of what we might call "classic", fully parallel OT by the start of the 2000s. The reader by McCarthy (2004b) presents a carefully edited selection of some of the most influential original papers from the "classic" period in the 1990s, also giving a good overview of the field. Much of this literature from the early OT period is available online through the Rutgers Optimality Archive (http://roa.rutgers.edu). Book-length, pedagogically oriented treatments are provided by Kager (1999) (an undergraduate-level textbook) and McCarthy (2008b) (perhaps more suitable for graduate-level study).

Many chapters in the handbook edited by de Lacy (2007) offer OT-focused overviews of several phonological subfields; in particular, Prince (2007) offers an introduction to the formal study of OT grammars *qua* theoretical objects. The paper by Vaux (2008), while highly critical of the OT enterprise, is highly useful in bringing together a large number of references to literature that intends to explicate the advantages of "classic OT". Scheer (2011) also provides a useful historical perspective on the development of OT within the broader phonological context.

Notes

1 Constraints can also be thought of as functions mapping structures to truth values (True or False, or “Violation” and “No Violation”): for a concrete proposal couched in model theory, see Potts & Pullum (2002). As Pullum (2013) points out, the OT architecture *as a whole* still remains procedural and thus largely incompatible with a model-theoretic approach to grammar. However, Heinz et al. (2009) discuss how it may be possible to provide a correct model of OT without explicitly resorting to the stepwise evaluation procedure.
2 *Gorgia toscana* is sometimes taken to refer only to the spirantization of voiceless stops; however, many relevant varieties also show spirantization of voiced stops as well, in common with other Italo-Romance varieties (Ramsammy, this volume).
3 Abbreviations: M = markedness (here *V[−son −cont]), F = faithfulness (here IDENT-IO), PF = positional faithfulness (here IDENT-IO_{gem}).
4 The more traditional approach focusing on inventories (cf. Hyman 2008) can also be accommodated: a gap in the inventory is created when the fully faithful candidate is defeated for the unattested input, so that input is mapped to something else.
5 I disregard here a very small number of borrowings, many of which have parallel forms conforming to the restriction.
6 For discussion of counterfeeding opacity, which *can* be accommodated in parallel OT, albeit at significant analytical cost, see McCarthy (2007a, sec. 2.3.3).
7 Potentially with some exceptions, as in the discussion of syllabification above.
8 I thank Joe Pater for discussion of this point.
9 Flynn (2007) provides some healthy skepticism on this point, however.

References

Alber, B., 2005. Clash, lapse and directionality. *Natural Language & Linguistic Theory*, 23(3), pp. 485–542.
Anderson, S.R., 1985. *Phonology in the twentieth century*, Chicago: University of Chicago Press.
Anttila, A., 1997. Deriving variation from grammar. In F. Hinskens, R. van Hout, & L. Wetzels, eds. *Variation, change and phonological theory*. Amsterdam: John Benjamins, pp. 35–68.
Archangeli, D., 1988. Aspects of underspecification theory. *Phonology*, 5, pp. 183–207.
Archangeli, D. & D. Pulleyblank, 1994. *Grounded phonology*, Cambridge, MA: The MIT Press.
Baković, E., 2007. A revised typology of opaque generalisations. *Phonology*, 24(2), pp. 217–259.
Benua, L., 1997. *Transderivational identity*. PhD thesis. Amherst: University of Massachusetts, Amherst.
Bermúdez-Otero, R., 2010. Morphologically conditioned phonetics? Not proven. Presentation OnLI II, University of Ulster. Available at: www.bermudez-otero.com/Belfast.pdf.
Bermúdez-Otero, R., 2011. Cyclicity. In M. van Oostendorp et al., eds. *The Blackwell companion to phonology*. Oxford: Blackwell Publishing, pp. 2019–2048.
Bermúdez-Otero, R., 2012. The architecture of grammar and the division of labour in exponence. In J. Trommer, ed. *The phonology and morphology of exponence: The state of the art*. Oxford studies in theoretical linguistics. Oxford: Oxford University Press, pp. 8–83.
Blaho, S., 2008. *The syntax of phonology: A radically substance-free approach*. PhD thesis. Tromsø: University of Tromsø.
Blumenfeld, L., 2006. *Constraints on phonological interactions*. PhD dissertation. Stanford: Stanford University.
Boersma, P. & S. Hamann, 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology*, 25(3), pp. 217–270.
Boersma, P. & B. Hayes, 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32(1), pp. 45–86.
Bye, P., 2001. *Virtual phonology: Multiple opacity and rule sandwiching in North Saami*. PhD thesis. Tromsø: University of Tromsø.

Calabrese, A., 2005. *Markedness and economy in a derivational model of phonology*, Berlin: Mouton de Gruyter.

Cassimjee, F. & C.W. Kisseberth, 1998. Optimal domains theory and Bantu tonology: A case study from Isixhosa and Shingazidja. In L.M. Hyman & C.W. Kisseberth, eds. *Theoretical aspects of Bantu tone*. CSLI lecture notes. Stanford, CA: CSLI, pp. 265–314.

Causley, T., 1999. *Complexity and markedness in Optimality Theory*. PhD thesis. Toronto: University of Toronto.

Chomsky, N. & M. Halle, 1968. *The sound pattern of English*, New York: Harper & Row.

Clements, G.N., 1986. Syllabification and epenthesis in the Barra dialect of Gaelic. In J.M. Stuart et al., eds. *The phonological representation of suprasegmentals*. Berlin: Walter de Gruyter, pp. 317–336.

Clements, G.N. & E.V. Hume, 1995. The internal organization of speech sounds. In J. Goldsmith, ed. *The Handbook of phonological theory*. Oxford: Blackwell, pp. 245–306.

Coetzee, A.W., 2008. Grammaticality and ungrammaticality in phonology. *Language*, 84(2), pp. 218–257.

Coleman, J., 1998. *Phonological representations: Their names, forms and powers*, Cambridge: Cambridge University Press.

De Lacy, P., 2006. *Markedness: Reduction and preservation in phonology*, Cambridge: Cambridge University Press.

De Lacy, P., ed., 2007. *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press.

Eisner, J., 1997. Efficient generation in primitive Optimality Theory. In *Proceedings of the eighth conference on European chapter of the association for computational linguistics*. EACL. Stroudsburg, PA: Association for Computational Linguistics, pp. 313–320.

Flemming, E., 2002. *Auditory representations in phonology*, London, New York: Routledge.

Flynn, D., 2007. Too many solutions: Says who? Presentation at 12th Alberta Conference on Linguistics. Available at: www.pec.ucalgary.ca/dflynn/files/dflynn/Flynn07b.pdf.

Frank, R. & G. Satta, 1998. Optimality Theory and the generative complexity of constraint violability. *Computational Linguistics*, 24(2), pp. 307–315.

Giannelli, L. & T.D. Cravens, 1996. Consonantal weakening. In M. Maiden & M. Parry, eds. *The dialects of Italy*. London, New York: Routledge, pp. 32–40.

Giannelli, L. & L.M. Savoia, 1978. L'indebolimento consonantico in Toscana. I. *Rivista italiana di dialettologia*, 2, pp. 23–58.

Giannelli, L. & L.M. Savoia, 1980. L'indebolimento consonantico in Toscana. II. *Rivista italiana di dialettologia*, 4, pp. 38–101.

Goldsmith, J.A., 1976. *Autosegmental phonology*. PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Gunnar Ólafur Hansson, 2001. *Theoretical and typological issues in consonant harmony*. PhD thesis. Berkeley: University of California, Berkeley.

Hale, M. & C. Reiss, 2008. *The phonological enterprise*, Oxford: Oxford University Press.

Hall, D.C., 2007. *The role and representation of contrast in phonological theory*. PhD thesis. Toronto: University of Toronto.

Halle, M., 1959. *The sound pattern of Russian: A linguistic and acoustical investigation*, 's-Gravenhage: Mouton & Co.

Hamann, S., 2009. The learner of a perception grammar as a source of sound change. In P. Boersma & S. Hamann, eds. *Phonology in perception*. Berlin: Mouton de Gruyter, pp. 111–149.

Hayes, B., 2000. Gradient well-formedness in Optimality Theory. In J. Dekkers, F. van der Leeuw, & J. van de Weijer, eds. *Optimality Theory: Phonology, syntax, and acquisition*. Oxford: Oxford University Press, pp. 88–120.

Hayes, B. & D. Steriade, 2004. Introduction: The phonetic bases of phonological markedness. In B. Hayes, D. Steriade, & R. Kirchner, eds. *Phonetically based phonology*. Cambridge: Cambridge University Press, pp. 1–33.

Hayes, B. & J. White, 2015. Saltation and the p-map. *Phonology*, 32(2), pp. 267–302.

Heinz, J., 2011a. Computational phonology part I: Foundations. *Language and Linguistics Compass*, 5(4), pp. 140–152.

Heinz, J., 2011b. Computational phonology part II: Grammars, learning, and the future. *Language and Linguistics Compass*, 5(4), pp. 153–168.
Heinz, J. & W.J. Idsardi, 2011. Sentence and word complexity. *Science*, 333(6040), pp. 295–297.
Heinz, J., G. Kobele & J. Riggle, 2009. Evaluating the complexity of Optimality Theory. *Linguistic Inquiry*, 40(2), pp. 277–288.
Honeybone, P., 2005. Sharing makes us stronger: Process inhibition and segmental structure. In P. Carr, J. Durand, & C.J. Ewen, eds. *Headhood, elements, specification, and contrastivity: Papers in honour of John Anderson*. Current Issues in Linguistic Theory. Amsterdam: John Benjamins, pp. 167–192.
Hyde, B., 2001. A restrictive theory of metrical stress. *Phonology*, 19, pp. 313–339.
Hyde, B., 2009. *A closer look at iterative foot optimization and the case against parallelism*. MS., Washington University of St Louis.
Hyman, L.M., 2008. Universals in phonology. *The Linguistic Review*, 25(1–2), pp. 83–137.
Inkelas, S., 1994. The consequences of optimization for underspecification. *NELS*, 27, pp. 287–302.
Inkelas, S., 1998. The theoretical status of morphologically conditioned phonology: A case study from dominance. In G. Booij & J. van Marle, eds. *Yearbook of morphology 1997*. Amsterdam: Springer, pp. 121–155.
Iosad, P., 2012. Final devoicing and vowel lengthening in Friulian: A representational approach. *Lingua*, 122(8), pp. 922–951.
Itô, J., A. Mester & J. Padgett, 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry*, 26, pp. 271–613.
Jurgec, P., 2010. *Feature spreading 2.0: A unified theory of assimilation*. PhD thesis. Tromsø: University of Tromsø.
Kager, R., 1999. *Optimality Theory*, Cambridge: Cambridge University Press.
Karttunen, L., 1998. The proper treatment of optimality in computational phonology. In *Proceedings of the international workshop on finite state methods in natural language processing*. Stroudsburg, PA: Association for Computational Linguistics, pp. 1–12.
Kenstowicz, M. & C. Kisseberth, 1979. *Generative phonology: Description and theory*, Orlando: Academic Press.
Kingston, J. & R.L. Diehl, 1994. Phonetic knowledge. *Language*, 70(3), pp. 419–454.
Kiparsky, P., 1982. Lexical phonology and morphology. In I.-S. Yang, ed. *Linguistics in the morning calm: Selected papers from SICOL-1981*. Seoul: Hanshin Publishing Company, pp. 3–91.
Kiparsky, P., 1985. Some consequences of Lexical Phonology. *Phonology Yearbook*, 2, pp. 85–138.
Kiparsky, P., 2000. Opacity and cyclicity. *The Linguistic Review*, 17(2–4), pp. 351–367.
Kiparsky, P., 2008. Universals constrain change: Change results in typological generalizations. In J. Good, ed. *Linguistic universals and language change*. Oxford: Oxford University Press, pp. 23–53.
Kirchner, R., 1998. *An effort-based approach to consonant lenition*. PhD thesis. Los Angeles: University of California.
Kirchner, R., 2000. Geminate inalterability and lenition. *Language*, 76(3), pp. 509–546.
Kisseberth, C., 1970. On the functional unity of phonological rules. *Linguistic Inquiry*, 1(3), pp. 291–306.
Köhnlein, B., 2016. Contrastive foot structure in Franconian tone accent systems. *Phonology*, 33(1), pp. 87–123.
Krämer, M., 2000. Voicing alternations and underlying representations: The case of Breton. *Lingua*, 110(9), pp. 639–663.
Krämer, M., 2012. *Underlying representations*, Cambridge: Cambridge University Press.
Lass, R., 1997. *Historical linguistics and language change*, Cambridge: Cambridge University Press.
Lombardi, L., 2001. Why place and voice are different: Constraint-specific alternations in Optimality Theory. In L. Lombardi, ed. *Segmental phonology in Optimality Theory: Constraints and representation*. Cambridge: Cambridge University Press, pp. 13–45.
Łubowicz, A., 2012. *The phonology of contrast*, London: Equinox.
Mackenzie, S., 2013. Laryngeal co-occurrence restrictions in Aymara: Contrastive representations and constraint interaction. *Phonology*, 30(2), pp. 297–345.

Magri, G., 2013. HG has no computational advantages over OT: Toward a new toolkit for computational OT. *Linguistic Inquiry*, 44(4), pp. 569–609.
Marantz, A., 1982. Re reduplication. *Linguistic Inquiry*, 13(3), pp. 435–482.
McCarthy, J.J., 1979. *Formal problems in Semitic phonology and morphology*. PhD thesis. MIT Press.
McCarthy, J.J., 1988. Feature geometry and dependency: A review. *Phonetica*, 45, pp. 84–108.
McCarthy, J.J., 1999. Sympathy and phonological opacity. *Phonology*, 16(3), pp. 331–399.
McCarthy, J.J., 2002. *A thematic guide to Optimality Theory*, Cambridge: Cambridge University Press.
McCarthy, J.J., 2003. Comparative markedness. *Theoretical Linguistics*, 29, pp. 1–51.
McCarthy, J.J., 2004a. Optimal paradigms. In L.J. Downing, A.T. Hall, & R. Raffelsiefen, eds. *Paradigms in phonological theory*. Oxford: Oxford University Press, pp. 170–210.
McCarthy, J.J., ed., 2004b. *Optimality Theory in phonology*, Oxford: Blackwell.
McCarthy, J.J., 2007a. *Hidden generalizations: Phonological opacity in Optimality Theory*, London: Equinox.
McCarthy, J.J., 2007b. What is Optimality Theory? *Language and Linguistics Compass*, 1(4), pp. 260–291.
McCarthy, J.J., 2008a. *Doing Optimality Theory: Applying theory to data*, Oxford: Blackwell.
McCarthy, J.J., 2008b. The gradual path to cluster simplification. *Phonology*, 25(2), pp. 271–319.
McCarthy, J.J., 2010. Autosegmental spreading in Optimality Theory. In J.A. Goldsmith, E. Hume, & W.L. Wetzels, eds. *Tones and features: Phonetic and phonological perspectives*. Berlin: De Gruyter, pp. 195–222.
McCarthy, J.J. & J. Pater, eds., 2016. *Harmonic grammar and harmonic serialism*, London: Equinox.
McCarthy, J.J. & A.S. Prince, 1994. The emergence of the unmarked: Optimality in prosodic morphology. *NELS*, 24, pp. 333–379.
McCarthy, J.J. & A.S. Prince, 1995a. Faithfulness and reduplicative identity. In J.N. Beckman, L. Walsh Dickey, & S. Urbanczyk, eds. *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*. University of Massachusetts, Amherst: GLSA, pp. 249–384.
McCarthy, J.J. & A.S. Prince, 1995b. Generalized alignment. In G. Booij & J. van der Marle, eds. *Morphology yearbook*. Dordrecht: Kluwer, pp. 79–153.
McCarthy, J.J. & A.S. Prince, 1999. Faithfulness and identity in prosodic morphology. In R. Kager, H. van der Hulst, & W. Zonneveld, eds. *The prosody–morphology interface*. Cambridge: Cambridge University Press, pp. 218–309.
Morén, B., 2001. *Distinctiveness, coercion, and sonority: A unified theory of weight*, London, New York: Routledge.
Moreton, E., 2004. Non-computable functions in Optimality Theory. In J.J. McCarthy, ed. *Optimality Theory in phonology*. Oxford: Blackwell, pp. 141–164.
Ní Chiosáin, M. & J. Padgett, 2001. Markedness, segment realization, and locality of spreading. In L. Lombardi, ed. *Segmental phonology in Optimality Theory: Constraints and representation*. Cambridge: Cambridge University Press, pp. 118–156.
Oostendorp, M. van, 2007. Derived environment effects and consistency of exponence. In S. Blaho, P. Bye, & M. Krämer, eds. *Freedom of analysis?* Studies in Generative Grammar. Berlin: Mouton de Gruyter, pp. 123–149.
Orgun, C.O., 1999. Sign-based morphology: A declarative theory of phonology-morphology interleaving. In M. van Oostendorp & B. Hermans, eds. *The derivational residue in phonological Optimality Theory*. Linguistik aktuell. Amsterdam: John Benjamins, pp. 247–267.
Orgun, C.O. & S. Inkelas, 2002. Reconsidering bracket erasure. In G. Booij & J. van Marle, eds. *Yearbook of morphology 2001*. Amsterdam: Springer, pp. 115–146.
Padgett, J., 2002. Feature classes in phonology. *Language*, 78(1), pp. 81–110.
Padgett, J., 2003. Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory*, 21, pp. 39–87.
Paradis, C., 1992. *Lexical phonology and morphology: The nominal classes in Fula*, New York, London: Garland.
Paster, M., 2012. Rethinking the "duplication problem". *Lingua*, 126(1), pp. 78–91.

Pater, J., 1999. Austronesian nasal substitution and other NÇ̥ effects. In H. van der Hulst, R. Kager, & W. Zonneveld, eds. *The prosody-morphology interface*. Cambridge: Cambridge University Press, pp. 310–343.
Pater, J., 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology*, 17(2), pp. 237–274.
Pater, J., 2009a. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In S. Parker, ed. *Phonological argumentation: Essays on evidence and motivation*. Advances in Optimality Theory. London: Equinox, pp. 123–154.
Pater, J., 2009b. Weighted constraints in generative linguistics. *Cognitive Science*, 33(6), pp. 999–1035.
Pater, J., 2016. Universal grammar with weighted constraints. In J.J. McCarthy & J. Pater, eds. *Harmonic grammar and harmonic serialism*. London: Equinox, pp. 1–46.
Potts, C., J. Pater, K. Jesney, R. Bhatt & M. Becker, 2010. Harmonic grammar with linear programming: From linear systems to linguistic typology. *Phonology*, 27(1), pp. 77–117.
Potts, C. & G.K. Pullum, 2002. Model theory and the content of OT constraints. *Phonology*, 19(3), pp. 361–393.
Prince, A.S., 2007. The pursuit of theory. In P. de Lacy, ed. *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, pp. 33–60.
Prince, A.S. & P. Smolensky, 1993. *Optimality Theory: Constraint interaction in generative grammar*. Technical report, Rutgers University Center for Cognitive Science; University of Colorado, Boulder.
Pulleyblank, D., 2006. Minimizing UG: Constraints upon constraints. WCCFL, 25, pp. 15–39.
Pullum, G.K., 2013. The central question in comparative syntactic metatheory. *Mind and Language*, 28(4), pp. 492–521.
Riggle, J., 2009. Violation semirings in Optimality Theory. *Research on Language and Computation*, 7(1), pp. 1–12.
Rose, S. & R. Walker, 2004. A typology of consonant agreement as correspondence. *Language*, 80(3), pp. 475–531.
Sagey, E., 1986. *The representation of features and relations in nonlinear phonology*. PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.
Scheer, T., 2010. *A guide to morphosyntax–phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*, Berlin: Mouton de Gruyter.
Scheer, T., 2011. Issues in the development of generative phonology. In N.C. Kula, B. Botma, & K. Nasukawa, eds. *The Continuum companion to phonology*. London: Continuum, pp. 396–446.
Schein, B. & D. Steriade, 1986. On geminates. *Linguistic Inquiry*, 17(4), pp. 691–744.
Schütze, C.T., 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*, Chicago: University of Chicago Press.
Scobbie, J.M., 1997. *Autosegmental representation in a declarative constraint-based framework*, New York: Garland.
Scobbie, J.M., J.S. Coleman & S. Bird, 1996. Key aspects of declarative phonology. In J. Durand & B. Laks, eds. *Current trends in phonology: Models and methods*. Salford: European Studies Research Institute, University of Salford, pp. 685–710.
Seeker, W. & D. Quernheim, 2009. Optimality Theory and vector semirings. In J. Piskorski, B. Watson, & A. Yli-Jyrä, eds. *Finite-state methods and natural language processing: Post-proceedings of the 7th international workshop FSMNLP 2008*. Frontiers in Artificial Intelligence and Applications. Amsterdam: IOS Press.
Shibatani, M., 1973. The role of surface phonetic constraints in generative phonology. *Language*, 49(1), pp. 87–106.
Sommerstein, A.H., 1974. On phonotactically motivated rules. *Journal of Linguistics*, 10(1), pp. 71–94.
Stanley, R., 1967. Redundancy rules in phonology. *Language*, 43(2), pp. 393–436.
Steriade, D., 1987. Redundant values. In A. Bosch, B. Need, & E. Schiller, eds. *Papers from the parasession on metrical and autosegmental phonology*. Chicago: Chicago Linguistics Society, pp. 339–362.

Steriade, D., 2001. *The phonology of perceptibility effects: The p-map and its consequences for prosodic organization*. MS., Massachusetts Institute of Technology.

Ternes, E., 1970. *Grammaire structurale du breton de l'Île de Groix (dialecte occidental)*, Heidelberg: Carl Winter Universitätsverlag.

Tesar, B., 2004. Using inconsistency detection to overcome structural ambiguity. *Linguistic Inquiry*, 35(2), pp. 219–253.

Tesar, B., 2013. *Output-driven phonology*, Cambridge: Cambridge University Press.

Tesar, B. & P. Smolensky, 2000. *Learnability in Optimality Theory*, Cambridge, MA: MIT Press.

Trommer, J., 2011. *Phonological aspects of Western Nilotic mutation morphology*. Habilitationsschrift, Leipzig: Universität Leipzig.

Trommer, J. & E. Zimmermann, 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology*, 31(3), pp. 463–510.

Vaux, B., 2003. Syllabification in Armenian, Universal Grammar and the lexicon. *Linguistic Inquiry*, 34(1), pp. 91–125.

Vaux, B., 2008. Why the phonological component must be serial and rule-based. In B. Vaux & A. Nevins, eds. *Rules, constraints and phonological phenomena*. Oxford: Oxford University Press, pp. 20–61.

Wareham, H.T., 1998. *Systematic parametrized complexity analysis in computational phonology*. PhD thesis. Victoria, BC: University of Victoria.

Wilson, C., 2001. Consonant cluster neutralisation and targeted constraints. *Phonology*, 18(1), pp. 147–198.

Youssef, I., 2015. Vocalic labialization in Baghdadi Arabic: Representation and computation. *Lingua*, 160, pp. 74–90.

3

Current issues and directions in Optimality Theory

Constraints and their interaction

Martin Krämer

3.1 Introduction

The first manuscripts on Optimality Theory (henceforth OT; Prince & Smolensky 1993/2004) were circulated in the early nineties, which is roughly a quarter of a century ago. For a theory, this is a long time, plenty of opportunity to develop, fracture, disintegrate or disappear into insignificance. As a quick look into the major journals of linguistics reveals, neither of the latter happened to OT. The majority of articles dealing with phonology in recent issues of the journals *Phonology*, *Linguistic Inquiry* and *Natural Language & Linguistic Theory* uses or discusses Optimality Theory. As far as *Linguistic Inquiry* is concerned one can even almost claim that phonology = OT, at least in the last four years. I found only one phonology-related article in *LI* in this period that didn't contain any reference to OT. The journal *Phonology*, however, provides a picture that is a bit more colourful, devoting substantial space to other theories/paradigms as well.

Seen from this perspective, we are thus dealing with the dominant paradigm in phonology these days. However, we will see in this article that the theory has fractured a bit into several competing models. There seems, however, to be a core of certain shared assumptions among the majority of scholars in the field – but there is also a vibrant debate and ongoing progress. To approach the question of current issues and directions in OT we will first have a quick look at the original basic assumptions to be able to better understand current debates (for a more detailed introduction, see Iosad, this volume). We will then look at which problems or challenges early OT caused or encountered and discuss which of these haven't been solved yet and which new problems have emerged with new developments. Given space constraints, the choice of topics that will be addressed here is necessarily limited and quite unlikely to be exhaustive (see as well Iosad, this volume, section 3, for discussion of some challenges to OT). Neither will it be possible to address any of these issues in great detail.

This chapter is structured as follows. Section 3.2 briefly recapitulates some of the basic tenets of OT to provide the background for the following discussions. Section 3.3 addresses the issue of the content or formalization of constraints, including functional grounding. Section 3.4 discusses constraint interaction. Here we will compare strict ranking with constraint weighting. The orthogonal matters of universal rankings versus more elaborate forms of constraint organization, such as stringency relations, and of constraint interaction by conjunction, i.e., the formation of complex constraints from more basic ones, will be discussed in section 3.5. The chapter ends with brief sketches of related topics in OT.

3.2 Basics

OT, as proposed by Prince & Smolensky (1993/2004), had a core of basic hypotheses that made it substantially different from other theories in phonology at the time. The basic idea was that grammatical output representations in a given language are optimal in comparison to conceivable alternative forms when evaluated against the language's idiosyncratic ranking of universal violable constraints on surface well-formedness (Markedness, henceforth M constraints) and on input–output mapping (Faithfulness, henceforth F constraints). A Generator function (Gen) computes a (potentially infinite) set of output candidates of which the Evaluation function chooses the one that is best according to the language-specific ranking of constraints. Thus, OT was conceived as a generative theory of constraint interaction, rather than a theory of constraints or representations or processes. This condensed summary contains a few hypotheses that we will look at in more detail now.

3.2.1 Parallelism

All candidates are available for comparison at the same time and all but one of them are filtered out simultaneously.

Traditional generative phonology (based on Chomsky & Halle 1968) is derivational. An input representation is assumed, and successively applied operations (rules) change this representation into different representations, the last one of which is the output of the derivation. In OT, one assumes an input and then compares potential output matches to this input to pick the best. There are no intermediate stages or representations. This was only one potential conception of OT. Prince & Smolensky entertained a serial evaluation function as well and then opted for the parallelist model. As McCarthy (1996, 1999, 2000, 2003, 2005, 2007 . . .) and many others pointed out, this caused a problem for the analysis of phonological opacity. Expressed in derivational terms, phonological opacity emerges when a later rule creates or removes the environment for a phonological process that has or was expected to apply earlier in the derivation (Kiparsky 1973). In more neutral terms, and under the default assumption that phonological processes apply whenever their conditions are present, we observe under- and overapplication, respectively. We can observe the result of the application of a process at the surface even though the environment for its application isn't present or we detect the environment for the application of a process, but it is blocked. There has been a range of proposals within OT to account for opacity or parts of the problem and related problems, such as Derived Environment Effects (DEE; or Non-Derived Environment Blocking, NDEB).

McCarthy (2007) discards all of them (including his own), some of them only on the grounds that they don't solve the whole problem "*en bloc*", and reintroduces serial derivation into OT with his Candidate Chains Theory (OT-CC), which subsequently developed into Harmonic Serialism (McCarthy & Pater 2016 and references there). A less radical return to serialism was proposed with Lexical Phonology & Morphology OT (LPM-OT) by Kiparsky (2000; see as well Rubach 1997), later branded as Stratal OT (Bermúdez-Otero 2011, forthcoming, this volume; Ramsammy, this volume). The extreme serialism of McCarthy's approach faces various problems of both a formal and empirical nature and developed back into Harmonic Serialism, a predecessor of OT. Another retrospective approach, van

Oostendorp's Coloured Containment (van Oostendorp 2004, 2007a, 2007b, 2017), retains parallelism but revives a repainted version of Prince & Smolensky's Containment model of Faithfulness. The original Containment model in OT was replaced by Correspondence Theory (McCarthy & Prince 1994, 1995, 1999), which can be considered the standard model of Faithfulness since the mid-nineties (see Iosad, this volume).

3.2.2 The Richness of the Base Hypothesis (RotB)

The heart of OT is summarized in the Richness of the Base Hypothesis, which states that all linguistic variation stems from the interaction of universal violable constraints rather than language-specific rules and restrictions on the lexicon. Since constraints are constraints on the output of evaluation and on the mapping between the input and the output, languages cannot differ by restrictions on the lexicon. Systematic differences between languages regarding their underlying representations have to be regarded as a side effect of the interaction of surface-oriented constraints. When we tease the RotB apart into sub-hypotheses we can isolate the following claims.

(1) Richness of the Base Hypothesis
  a) Surface patterns of languages are determined by constraints.
  b) Constraints are potentially conflicting, and therefore violable and rankable.
  c) Constraints are universal.

Despite McCarthy's (1993) proposal of a language-specific rule of r-insertion in Gen for an analysis of English intrusive *r*, sub-hypothesis (a) hasn't been challenged among proponents and users of OT. Also sub-hypothesis (b) enjoys widespread acceptance, though see, e.g., Orgun & Sprouse (2010) for the proposal of inviolable constraints to model absolute ungrammaticality. Whether constraints are universal, though, is a constant matter of debate in one form or other. As we will see in section 3.3, whether constraints are universal or not also depends, at least for some constraints, on our interpretation of the term "universal" (for a non-universalist, emergentist viewpoint, see Archangeli & Pulleyblank 2015, this volume).

3.2.3 Constraints and their properties

OT didn't come with a fully articulated theory of constraints. The few things we can say about constraints are that they are either M or F constraints, that they are violable, potentially conflicting, statements about output representations and about the mapping between correspondent representations at different levels (e.g., input and output – see Krämer 2012 for a discussion of the terms input and output in the context of OT) and that they are arranged in language-specific dominance hierarchies. Constraints that are proposed by an analyst to account for a certain phenomenon should be motivated, or grounded, either typologically or functionally, ideally in both ways. Typologically grounded constraints show dominance, i.e., a direct effect – rather than a residual effect – in some language. A constraint whose effect is only observed in one language and there only indirectly, i.e., the constraint is dominated by other conflicting constraints, is thus not a typologically well-motivated constraint. The poster children of typologically well-grounded constraints are Onset and *Coda. Many languages don't allow syllables to start in a vowel and many languages don't display syllables closed by a consonant. In such languages these constraints are undominated. In other languages, though, they show residual effects. Intervocalic consonants are usually syllabified as syllable onsets (rather than codas) even in languages that allow onsetless syllables and closed syllables.

Constraints are functionally grounded if they are motivated by some observation about articulation, aerodynamics or perception (see Ramsammy's contribution on the phonetics interface in OT, this volume, and, for example, Gordon's 2007 overview of functionalism in phonology). Certain segments are more difficult to articulate than others for reasons to be found in airstream mechanisms or in the vocal tract. Thus, a constraint against voiced obstruents, e.g., the Voiced Obstruent Prohibition (VOP, *[+voice], *Laryngeal) can be said to be functionally grounded, since vocal fold vibration is difficult to sustain if the outgoing airstream is blocked. Regarding the claim to universality, one could reasonably argue that functionally grounded constraints are good candidates for universal constraints since all humans share the same physical and mental capacities and limitations. However, being grounded this way, these constraints don't need to be hard-wired into the genome (i.e., not part of Universal Grammar), since they could be "discovered" by every individual through induction (see Hayes 1999 on inductive grounding).

However, constraints that are only typologically grounded but crucially lack a functional motivation can only be innate – in the sense that they are part of the genetically determined part of the language faculty. We will come back to this issue in section 3.3.

Discussing the formalization and properties of constraints in OT is a complex matter also because OT doesn't come with a theory of phonological representations. For example, while the majority of constraints referring to phonological features in Prince & Smolensky (1993/2004) and McCarthy & Prince (1995) makes reference to binary valued features (especially IO-Identity[±F] constraints), nothing hinges on this choice. OT has successfully been used with privative features and elements and even been combined with Exemplar Theory (see van de Weijer 2012 for the latter). However, the primitives the constraints refer to can be expected to make a difference for their formalization and their systemic properties.

However, there are some properties each OT constraint should have. One of these is a clear definition of its violation profile, i.e., for any constraint it should be unambiguous which structures violate it and which satisfy it. And the violations of each constraint should be comparable to those of every other constraint. A property concerning violability that constraints differ on is categoriality. Most constraint violations are categorical. Either a structure violates a constraint or it doesn't. In addition we can count the number of violations, i.e., we can quantify how bad a candidate fares with respect to a constraint and compare this with any other candidate's performance on the same constraint or any other. However, the only thing that matters is whether a candidate has the same or a different number of violations than a candidate it is compared with. "Quantification" thus necessitates counting only up to one. Constraints are gradiently violable if a single constraint violation can be more or less severe. This was considered a problem with Alignment constraints (alignment: McCarthy & Prince 1993; problem with non-categoriality of alignment: McCarthy 2003, Hyde 2012 and references cited there). While there is a categorical interpretation of Alignment constraints, their violations have usually been assessed gradiently. One might think that it is an either-or question whether two edges coincide or not; however, the general practice established by McCarthy & Prince (1993) is that Alignment constraints compute the distance between two edges in a candidate by referring to a third, the intervening, category. This was very useful in most cases, but also ran into the problem referred to as the "midpoint pathology" (an Alignment constraint places stress on the only foot at the centre of a string rather than at an edge) among other things.

3.2.4 Constraint interaction

As sketched above, constraint interaction is the source of cross-linguistic variation. The basic form of constraint interaction is considered to be ranking in a hierarchy. These dominance relations were hypothesized to be strict and transitive. They are strict in the sense that if constraint *A* dominates constraint *B*, candidate *a*, preferred by constraint *A* over candidate *b*, is considered more harmonic than candidate *b*, regardless of the number of violations candidate *a* incurs on constraint *B* in excess of those of candidate *b*. If candidate *a* satisfies constraint *A* and violates constraint *B* many times, while candidate *b* violates constraint *A* only once and even satisfies constraint *B*, the former candidate is still more harmonic than the latter, as illustrated in the tableau in (2). (In OT tableaux, each asterisk indicates a constraint violation and the pointing finger the preferred candidate. Constraints separated by a full line are ranked with respect to each other, with the constraint to the left dominating the constraint to its right.)

(2) Strictness of strict domination

	A	B
☞ a		***********
b	*!	

Transitivity can be observed in ranking arguments. If constraint *A* can be argued to dominate constraint *B* and constraint *B* dominates constraint *C*, then constraint *C* is also dominated by constraint *A* by transitivity. No evidence/ranking argument is needed to establish the relation between constraint *A* and constraint *C*. If constraint *C* has to be assumed to dominate constraint *A* we are facing one type of ranking paradox. The most straightforward ranking paradox emerges when one piece of data requires *A* dominating *B* while another piece of data in the same language requires the reverse ranking of the same constraints.

Constraints could also be ranked in a non-strict way. In such a model, every constraint carries a certain weight, and violations of a constraint with lower weight could add up to outweigh the fewer violations of a constraint with more weight incurred by a competing candidate. Thus, the constraint with less weight potentially reverses the decision made by a constraint with more weight. Alternatively, rankings could be more flexible, with mobile constraints "inhabiting" potentially overlapping zones on a ranking scale. The latter two types of ranking have been proposed in Harmonic Grammar (Legendre, Miyata & Smolensky 1990; Goldsmith 1990; Pater 2016) and Stochastic OT (Boersma 1998 et seq.), respectively. We will discuss the former, weighted constraints, in section 3.4.

Transitivity can potentially be circumvented, or relativized, if lower ranked constraints are allowed to team up against higher ranked constraints or if they can be cloned, with the resultant complex constraints or clones inhabiting higher strata in the hierarchy than the otherwise dominating conflicting constraints. After considering these options, local conjunction and positionally restricted constraints (positional Faithfulness, positional Markedness) as well as constraint indexation, we will look at other modes of constraint interaction, universal rankings versus stringently organized constraints in section 3.5.

3.3 The motivation and formalization of constraints

3.3.1 Grounding

As sketched in the previous section, OT constraints are expected to be grounded either typologically or functionally, rather than just stipulated or descriptively convenient or a rephrasing of a generalization or rule. There has been a considerable body of work investigating how the acoustic signal and the limits and biases of human perception shape phonological patterns and phonological computation (e.g., contributions in Hayes, Kirchner & Steriade 2008).

The hypothesis that phonology is grounded in phonetics could as easily be reversed into phonological grounding, that is, phonetics is grounded in phonology, since for the arguments brought up by functionalists we often can't tell which is the cause and which is the effect. Liljencrants & Lindblom's (1972) observation on the dispersion of vowel systems is a typical example. Why are vowels dispersed in the vowel space? The functionalist's answer is that this is the case to make perception easier. The thing is that what is made easier is the perception of phonological contrasts. If two phonetic exponents of phonological objects are intended to encode the contrast between the two they should be perceptually distinguishable. Contrast is a, if not *the*, phonological function. Thus, the driving force here is not the phonetics (or perceptual difficulties) but the phonology.

Coming from this perspective, Krämer (2012), for example, argues for replacing phonetically defined features by a set of abstract contrastive features loaned from other cognitive modules (syntax, semantics), which get mapped to an articulation or signal that best corresponds to the semantics of the respective feature. Thus, the phonological content determines the phonetic shape of its exponent.

One can extend this kind of reasoning to the grounding of OT constraints. The prevalent greater diversity of segmental contrasts in Onset or presonorant position has been attributed to the availability of better phonetic cues for contrasts, e.g., voicing in obstruents (Steriade 2009). Steriade proposes the hierarchy of contexts favouring/disfavouring voicing contrast in obstruents given in (3). This contrast is found either nowhere in a language or only in presonorant position or in presonorant position and word-finally, or in these positions and after sonorants and so on.

(3) Steriade's hierarchy of favourable environments for obstruent voicing contrast
nowhere – presonorant – word-finally – postsonorant – preobstruent – everywhere

Positionally restricted contrast can be analyzed by recourse to positional Faithfulness constraints. In this situation the choice is between structurally determined abstract constraints, limiting the scope of the constraint by syllable position, i.e., Faithfulness to onsets and other positions, such as edges of certain domains, the morpheme or the Prosodic Word (e.g., Faith(F)/edgeX) or to define the constraints by their phonetic environment, i.e., Faith(F)/presonorant. The latter approach runs into a minor empirical problem when the respective phonetic contexts are given but some more abstract structure makes them invisible or irrelevant for phonological computation. A language might have a certain contrast in obstruents in presonorant position, but only if both the preceding obstruent and the sonorant fulfil additional conditions. Such conditions could be of morphological nature, i.e., that they belong to the same morpheme or word; or they could be of phonological nature, i.e., that the two involved segments belong to the same syllable or foot. Thus, languages with a voicing contrast might have final devoicing, but still devoice obstruents in some presonorant positions because the following sonorant is not in the same word, morpheme, syllable, foot or phrase as the obstruent.

Northern varieties of German, for example, display final devoicing in presonorant position only under certain conditions. The voiced stops devoice before tauto-morphemic nasals, but not before tauto-morphemic laterals. While a word like *Flug* 'flight' has a final [k] or [g̊], the plural form *Flüge* shows the underlying /g/. We find the same alternation in pairs like *regnen* [ʀeːknən] 'to rain', which displays the voiceless or neutralized dorsal in presonorant position, and *Regen* [ʀeːgən] 'rain', which has a voiced dorsal stop. In an only slightly different context, though still in presonorant position, namely before a lateral, as in *Regler* 'modulator', /g/ surfaces with its underlying voicing specification, though it doesn't in *behaglich* 'comfortable'. In the latter case the morpheme boundary is between the obstruent and the lateral, while in the former it follows the lateral (i.e., *behag-lich* versus *Regl-er*).

The situation becomes even more clear in varieties that also have g-spirantization. In such varieties we get *Flu*[x] – *Flü*[g]*e* and *Re*[g]*en* 'rain' – *re*[ç]*nen* '(to) rain' but Re[g]*el* 'rule' – *Re*[g]*ler* 'ruler, modulator'.[1]

In the licensing by cue approach we would expect **re*[g]*nen* (which is licit in some varieties) and **beha*[g]*lich* or, alternatively, **Re*[k]*ler*. It is difficult to imagine an analysis that doesn't make reference to syllable constituents. /g/ is devoiced and spirantized in the coda, and in words such as *regnen*, the *gn* sequence is syllabified in two syllables, i.e., [ʀeːç.nən], rather than *[ʀeː.gnən], whereas *gl* in *Regler* is an Onset, [ʀeː.glɐ], as in words like *Glück* 'happiness/luck'. While /gn/ is a licit syllable Onset, as in *Gnom* 'gnome', this type of Onset is very infrequent and seems to be marked (compare the historical fate of English stop + nasal onsets). Even without consulting a dictionary or frequency database one can safely say that words starting in *gl* are fairly unmarked.

However, the same generalizations on voicing neutralization hold for labial and coronal stops, and there are no words in German that start in the sequence /dl/. Surprisingly (for the syllable-based account, but not for the licensing by cue approach), word- or morpheme-internally /d/ doesn't devoice before a lateral, as in *Adler* 'eagle'.

If one considers place features (Place of Articulation, henceforth PoA), one could easily come up with the conclusion that PoA in stops is better cued in postvocalic or postsonorant position, since we find the cues to the PoA of the stop in the formants of the vocoid in the transition from the vocoid to the closure of the stop. In the release phase, on the other hand, there can be the laryngeal burst of aspiration or the delayed VOT masking the transition of articulators from the stop's PoA to that of the following vowel or sonorant consonant. In addition, that following sonorant consonant could have only weak formant structure compared to a vowel, which provides bad cuing. Also, there often is no release, i.e., at the end of a word/phrase/utterance or before another obstruent.

A typological observation that corroborates the idea of preemptive cuing (see as well Kochetov & So 2007) is nasal place assimilation (NPA). NPA is usually regressive (while **co*[n]*position* and **a*[n]*chor* don't sound good in English, *apnoea* and *acknowledge* are fine). Postnasal stops tend to expand their PoA into preceding nasals. Progressive NPA, as in colloquial northern German (e.g., [ʔamp] *Amt* 'office' or [habm̩] *haben* 'to have'), is much more exotic. This could be functionally grounded by concluding that, in comparison to a vowel, the transition from a nasal with its relatively weak formant structure provides a bad cuing ground for the place properties of a following stop. Hence, to increase perceptibility, the PoA of the stop is extended over the whole duration of the preceding nasal. This reasoning then also leads to the conclusion that word-initial position is a bad place for PoA contrasts in stops/obstruents, since the stop starts with silence and the cues for the PoA emerge late and might be masked. Thus, PoA contrasts should typologically be more common in postvocalic position than in postsonorant position than in word-initial position. As it happens, PoA of obstruents is typically neutralized in coda, or preobstruent and pre-pause, position. In addition, many languages that have word-internal neutralization of contrasts in coda/preconsonantal position do not neutralize PoA in word-final consonants, even though stops in this position often don't even have a release (see, e.g., some of the cases in Piggott 1999). West Greenlandic neutralizes PoA in word-initial fricatives by allowing only the alveolar or the laryngeal fricative, while the language displays fricatives at seven distinct PoA in other positions (Rischel 1974; Fortescue 1984). This pattern is surprising for two reasons. First, one would think that fricatives have intrinsic PoA cues and show PoA contrasts in more contexts than stops. Second, word-initial position is (in this language) necessarily presonorant/prevocalic. If, then, (post-pausal) prevocalic PoA is less well-cued than PoA in postvocalic, or intervocalic, position, the common tendency to display more PoA contrasts in word-initial or generally syllable Onset position has to stem from another cause than licensing by cue. This other cause could be the nature of Faithfulness or Licensing constraints, i.e., the phonology.

Functional/phonetic grounding feeds the innateness debate. "Good" M constraints are those that are grounded in articulatory difficulty or complexity or in limits of perception etc. Since all human bodies are by and large the same in all relevant respects, one could say that M constraints then don't have to be universal, in the sense of innate (hard-wired in the genome), since every language learner can infer the constraints from the data, either by monitoring her own production (problems) or her own perception – though the latter must be more of a challenge, given that learners tend to ignore overt correction, i.e., negative feedback.

In some of the functionally leaning literature constraints are therefore language-specific, learned in first language acquisition through overt evidence by exposure to the ambient language (Boersma 1998; Hayes 1999; Hayes & Wilson 2008). This stance raises two questions, one of which is how constraints are learned that are responsible for static phonotactic restrictions, the other is how constraints have been learned that only show effects in interlanguage, i.e., when adults learn a second language.

For simple M and related F constraints the situation is quite straightforward. A learner discovers a contrast, say, a minimal pair, such as *back* vs. *pack* in English, and infers that there must be an M constraint against the articulatory more marked member of the opposition, i.e., *[spread glottis] or *[voice] (depending on your favoured analysis of English), and a conflicting F constraint, e.g., IO-Ident(laryngeal). The situation already gets difficult if the pattern involves positional neutralization. There are functional argumentations for both positional Faithfulness as well as positional Licensing/Markedness. Which choice does a learner make? Maybe she needs both, as one would for Dutch voice patterns (Grijzenhout & Krämer 2000), but if only one is needed, as for German, which displays final devoicing but no voicing assimilation, there is no way of choosing (see also the discussion of positional restriction of constraints in the next section).

When it comes to learning static phonotactic restrictions a learner actually needs negative evidence to induce constraints, which she doesn't get (see Prince & Tesar 2004, but also Hayes & Wilson 2008 for phonotactic learning in OT). Consider for example Beijing Chinese, which allows only nasals in postvocalic position (Blevins 1995), has only mono-syllabic words and hardly any morphology, and thus no relevant alternations. A learner who is expected to learn the Coda Condition of Beijing Chinese has a problem. She could figure out that there are perceptual and articulatory challenges in postvocalic/preconsonantal/pre-pausal obstruents and liquids if she had to perceive or produce any. Since she is never confronted with any challenge of this sort, she has no way of figuring out that there is a constraint that bans the liquid and the obstruents from the coda.

This result is challenged by data from L2 acquisition (Broselow, Chen & Wang 1998). Chinese learners of English show strategies to avoid coda consonants, in words such as *bag* or *hammock*, that range from featural changes to deletion of the offensive coda consonants, or epenthesis of a vowel. The latter moves the consonant into Onset position. First, such L2 processes can be seen as an argument for the superiority of constraint-based phonology over rule-based theories, since it is more plausible that such processes, such as devoicing of voiced coda consonants, can be explained straightforwardly as effects of constraints that are present in the Chinese grammar, but never show an effect since the Chinese lexicon doesn't contain any items that would create the context for the constraints to exert an influence on the choice of output candidate. Second, and more importantly for the discussion here, this implies that such constraints are present in the Chinese grammar, which is unexpected from the purely emergentist approach. The short response to this challenge is to assume that all constraints and representational building blocks (e.g., features or feet) are universal in the sense of being innate, i.e., hard-wired into the human genome (as assumed by Prince & Smolensky 1993/2004).

The less bold and slightly more complex stance refines the emergentist view: Constraints are universal in the sense that every human mind exposed to spoken (or signed) language can and does draw the same conclusions about representational options and about markedness (Hayes & Steriade 2004, see also Collins 2013 for a more refined discussion, and Archangeli & Pulleyblank, this volume).

3.3.2 Positional restrictions and the definition of constraints

A discussion completely different from the universality question, but strongly connected to the grounding issue, has been simmering for years. There are several proposals for the analysis of positional neutralization. The dispute boils down to the question whether positional neutralization should be regarded as a certain contrast being allowed only in a certain position or whether it is banned from the complementary position. The competing approaches, i.e., positional Faithfulness, positional Licensing and positional Markedness, are all functionally grounded (claiming better or worse perceptual/articulatory conditions in the respective complementary environments).

The problem can easily be illustrated with final devoicing, already alluded to above. Many languages display a voicing or other laryngeal contrast only in obstruents in the syllable Onset or presonorant position. Often this interacts with voicing assimilation.

Positional Faithfulness (Beckman 1997, 2004; Lombardi 1999) assumes a clone of general Faithfulness to the laryngeal feature, restricted to obstruents in Onset position, IdentityOnsetLar, which interacts with a simple M constraint, as illustrated in tableau (a) in (4). In most of the literature that uses final devoicing as a pet example to make some theoretical point, but which isn't interested in the typology of voicing patterns, a positional Markedness constraint is assumed that is violated by voiced obstruents in coda position, as in tableau (b) in (4) (see, e.g., Pater's 2016 arguments against Local Constraint Conjunction, as well the discussion below in section 3.5).

(4) Positional Faithfulness or positional Markedness

a.	/bad/	IdentOnset	*Lar	Ident
i.	bad		**!	
ii.	pat	*!		**
iii.	pad	*!	*	*
☞ iv.	bat		*	*

b.	*Lar/coda	Ident	*Lar
	*!		*
		**!	
	*!	*	**
		*	*

Consideration of voicing assimilation, which is usually regressive if the two members of an obstruent cluster are in separate syllables, settles the issue in favour of the positional Faithfulness analysis. The positional Markedness analysis erroneously predicts progressive devoicing rather than regressive voicing for inputs with a voiced obstruent on the right. At this point one could also dismiss positional Licensing (Zoll 2004; Walker 2011), since positional Faithfulness together with simple Markedness does the job. Consider the next tableau, in which I added a Licensing constraint, which demands that laryngeal features be licensed by or linked to a segment in an Onset.

(5) Positional Faithfulness and positional Licensing

	/bad/	IdentOnset	*Lar	Ident	Lic(lar)/Onset	*Lar/coda
i.	bad		**!		*	*
ii.	pat	*!		**		
iii.	pad	*!	*	*	*	*
☞ iv.	bat		*	*		

At first sight, the Licensing constraint and the positional Markedness constraint seem indistinguishable. If we consider an assimilation situation in a language with final devoicing it becomes obvious that the Licensing approach is compatible with the typological observation when combined with positional Faithfulness. A voiced stop in non-onset position (i.e., coda) that shares the feature with a stop in Onset position satisfies the Licensing constraint, but still violates the positional Markedness constraint. However, the Licensing constraint can't explain the directionality of assimilation. It needs positional Faithfulness.

(6) Positional Faithfulness and positional Licensing in regressive assimilation

	/. . . pd . . ./	Agree	IdentOnset	*Lar	Ident	Lic(lar)/Onset	*Lar/coda
☞ i.	b.d			**!	*		*
ii.	p.t		*!		*		
iii.	p.d	*!		*			

However, the situation is more intricate. Zoll (2004) provides examples in which also derived marked structures are allowed in certain prominent domains only, rather than only contrastive, i.e., underlying, marked features. This kind of positional restriction can't be captured by positional Faithfulness alone, since positional Faithfulness only caters for marked structures that are present in the input already. Walker (2011) argues that processes in which the trigger is in a weak position and the target in a prominent position (unstressed syllable and stressed syllable, respectively) are an effect of licensing. The contrastive feature has to be licensed by a prominent position.

On the other hand, positional Licensing doesn't account for directionality effects in assimilation processes, as we have just seen. For example, the difference between stress- or stem-controlled vowel harmony and metaphony, if assumed to be caused by Licensing constraints, has to be attributed to positional Faithfulness constraints. In vowel harmony, all vowels in non-prominent positions assimilate to the vowel in the initial syllable, the stressed syllable or the next syllable in the stem (see, e.g., Krämer 2003a for an overview). In metaphony, on the other hand, a word-final or unstressed vowel causes the vowel in the stressed syllable to assimilate (see, e.g., Walker 2011 or the contributions in Torres-Tamarit, Linke & van Oostendorp 2016). Vowel harmony systems require a high-ranking Faithfulness constraint, restricted to a prominent position. Metaphony requires a highly ranked positional Faithfulness constraint restricted to the last syllable (Walker 2011; see Krämer 2003a, 2003b for a discussion of this type of edge Faithfulness). In the following tableaux, this constraint interaction is illustrated. Rearranging the two positional Faithfulness constraints yields either the metaphony candidate (c), as in (7), or the vowel harmony candidate (b), as in (8), as optimal.

(7) Metaphony as licensing

–F +F /CV'CVCVCV/	LIC(F)/P	IDENT(F)/R	IDENT(F)/Px	IDENT(F)
a. –F +F CV'CVCVCV	*!			
b. –F CV'CVCVCV		*!		***!
☞ c. +F CV'CVCVCV			*	**!
d. +F CV'CVCVCV			*	***

(8) Harmony as licensing

–F +F /CV'CVCVCV/	LIC(F)/P	IDENT(F)/Px	IDENT(F)/R	IDENT(F)
a. –F +F CV'CVCVCV	*!			
☞ b. –F CV'CVCVCV			*	***
c. +F CV'CVCVCV		*!		**
d. +F CV'CVCVCV		*!		***

The individual approaches undergenerate by not producing attested patterns and overgenerate certain unattested patterns. Admitting all options, i.e., positional Faithfulness, positional Licensing and positional Markedness, accounts for attested patterns, but also produces undesired/unattested patterns.

Assimilation patterns, such as voicing assimilation, vowel harmony or metaphony, have been subject to analyses with a range of different constraints as the cause for the patterns – Agree constraints, Syntagmatic Correspondence, ABC theory, positional and simple M, positional Licensing, different flavours of Alignment constraints (see Beckman 1997; Lombardi 1999; Baković 2000; Krämer 2003a; Walker 2005, 2011; Jurgec 2011) – and it looks as if the issue is still far from settled.

The most appealing approach is of course the one that doesn't need a constraint that is only postulated to account for assimilation. Beckman (1997) attempts this. In her analysis of vowel harmony, only positional Faithfulness and simple Markedness constraints generate the pattern she discusses. The approach does not only run into the problems raised above, it also opens for another question: Should M constraints also exist for the unmarked value of a feature? The problem is illustrated in the following tableau, which schematizes stress-controlled vowel harmony. Without the constraint referring to the unmarked value, candidates (a) and (b) can't be distinguished.

(9) Unmarkedness constraint

–F +F –F –F | | | | /CVˈCVCVCV/	Ident(F)/Px	*+F	*–F
–F +F –F –F | | | | a. CVˈCVCVCV		*	*!**
+F (linked to all vowels) ☞ b. CVˈCVCVCV		*	
–F –F –F –F | | | | c. CVˈCVCVCV	*!		****
–F (linked to all vowels) d. CVˈCVCVCV	*!		*

Parsimony demands that M constraints punishing unmarked structure should not be included in the constraint set, since they double the set of simple M constraints (at least those referring to features), and since the assumption of their existence requires a fixed ranking between M constraints referring to the marked value (*+F) and the corresponding M constraint referring to the unmarked value (*−F), which is *+F dominating *−F. The latter undermines the Free Ranking Hypothesis, which postulates that all constraints can be ranked freely (Prince & Smolensky 1993/2004). The Free Ranking Hypothesis was challenged already by Prince & Smolensky themselves in their discussion of harmonic alignment, sound patterns which apparently require universal rankings of constraints that refer to a scale, such as the sonority hierarchy. The problem of fixed hierarchies will be taken up again in section 3.5.

Scalar well-formedness brings us to our next issue to be considered here, the nature of violability as binary. Constraint violation could be binary (or categorical), i.e., a statement on well-formedness is true for a certain representation or it isn't, or violation could have numerical or other scalar values, i.e., a representation occupies a certain rank on a scale of fitness. In the latter interpretation, constraints could be mildly violated by one representation and severely violated by another.

Prince & Smolensky (1993/2004) discuss syllable positions in this respect. An ONSET is more harmonic the lower it is on the sonority hierarchy, while a nucleus is the more harmonic the higher it is on the sonority hierarchy. Thus, M constraints on these positions could assess a syllable position's value with regard to its distance from the highest or lowest class on the sonority hierarchy.

(10) Onset Harmony (H-ONS): The left edge of a syllable is more harmonic the lower it is on the sonority scale. Assign one violation mark for every step on the harmony scale an ONSET is away from the sonority of stops.

Table 3.1 Sonority and cumulative violation

	vowel	liquid	nasal	fricative	stop	
	****	***	**	*	√	

In comparison, a constraint such as ONSET, which requires every syllable to start with a consonantal ONSET, can also be violated to various degrees by a single candidate, but then it is violated in different locations, i.e., by several syllables. Every syllable itself either satisfies or violates the constraint.

(11) ONSET: Assign a violation mark for every syllable that does not have a consonantal ONSET.

Table 3.2 Gradience versus categoriality

		ONSET	H-ONS
a.	.æ.æ.æ.	1+1+1	4+4+4
b.	.æ.æ.	1+1	4+4
c.	.æ.	1	4
d.	.læ.		3
e.	.næ.		2
f.	.hæ.		1
g.	.ʔæ.		0

Languages use such scales in various ways, however, that are problematic for OT. One problem is that languages conflate or telescope the levels of such hierarchies, and the other is that they use seemingly random points on such scales as tolerance thresholds. E.g., while most languages tolerate only vowels as syllable nuclei, some also allow sonorant consonants and some display obstruents in nucleus position. However, with strict ranking, a constraint like H-NUC (the sister constraint of H-ONS, defining nucleus well-formedness) either dominates Faithfulness or it is dominated by F constraints. It thus doesn't matter how bad a nucleus or ONSET is according to the respective scalar constraint unless each of these constraints is decomposed into binary sub-constraints that are rankable with respect to F constraints.

(12) Scalar constraints decomposed I
H-ONS = {*[+vocalic]/Ons ›› *[+liquid]/Ons ›› *[+son]/Ons}

(13) Gradience versus categoriality

		H-Ons	*[+voc]/Ons	*[+liquid]/Ons	*[+son]/Ons
c.	.æ.	4	*		*
d.	.læ.	3	*	*	*
e.	.næ.	2			*
f.	.ʔæ.	1			

To keep the implicational result of the scalar constraint, these constraint sets have to be in some kind of universal relation. Prince & Smolensky (1993/2004) propose a universal ranking. This undermines the Free Ranking Hypothesis, the assumption that all constraints can potentially be ranked in any order. We will take up other solutions to this problem in section 3.5.

A similar scalarity issue emerges with Alignment constraints (McCarthy 2003). Even though the definition requires mapping of two edges, as can be read from the definition in (14), the constraints are usually used to actually measure the difference between two edges by way of some intervening category. Thus the constraint violations are computed according to the clause in (15) (see discussion in McCarthy 2003; Hyde 2012).

(14) Generalized Alignment (McCarthy & Prince 1993)
Align (Cat1, Edge1, Cat2, Edge2)
The Edge1 of every Cat1 coincides with the Edge2 of some Cat2.

(15) Generalized Alignment, the third argument (McCarthy 2003)
Assess a violation mark for every Cat3 that intervenes between edges that fail to coincide.

To illustrate this we consider a constraint that requires all feet to be at the right edge of a word. (16) and (17) provide the usual gradient and a categorical definition, respectively. The tableau in (19) shows violation profiles for the two hypothesized constraints in a grammar that doesn't allow proper edge mapping by way of a higher ranked constraint, Non-Finality, which keeps the last syllable in the word extrametrical.

(16) AllFtR (usual version): Align(Foot, R, Wd, R) Align the right edge of every foot with the right edge of a Prosodic Word. Assign a violation mark for every syllable between each foot and the right edge of the word.

(17) ∀FtR (oversimplified categorical version): Assign a violation mark for every foot that is not at the right edge of the word.

(18) Non-Finality: The rightmost syllable in a Prosodic Word is not parsed in a foot.

(19) Edge magnetism and categoriality

		Non-Final	∀FtR	AllFtR
a.	{σσσσ(σσ)}	*		
∀/A b.	{σσσ(σσ)σ}		*	*
∀ c.	{σσ(σσ)σσ}		*	**
∀ d.	{(σσ)σσσσ}		*	***
e.	{σ(σσ)(σσ)σ}		*, *	***, *
f.	{(σσ)σ(σσ)σ}		*, *	****, *

The categorical version of the Alignment constraint doesn't distinguish candidates (b), (c) and (d). However, in real-world cases like this, feet usually huddle up at the designated edge even if, due to a higher ranked constraint, alignment can't be perfect – compare candidates (e) and (f). Thus, the gradience of Alignment produces actually a desired result. It is other scalar constraints that are pointless or counter-productive in OT (see the discussion earlier in this section, as well as below in section 3.4).

Eisner (1997) observes the midpoint pathology produced by gradient Alignment constraints. Under certain circumstances, an Alignment constraint can drag a structure to the centre of a domain. Since phonological processes and structure building are usually edge oriented this is an undesired result. Feet and stress are usually oriented towards the left or right edge of the word, not the centre.

This midpoint pathology is illustrated in (20). The Alignment constraint requires the left edge of every syllable to align with the left edge of some foot. For some reason (higher ranked constraint) only one foot is allowed per word. Violation marks contributed by individual syllables are separated by commas, while violation marks associated with syllables preceding the foot, syllables inside the foot and syllables following the foot respectively are separated by semicolons. On the right I have given the total violation score (Σ) of each candidate for convenience, followed by violations incurred by syllables preceding the foot (p), contained within the foot (c) and following the foot (f).

(20) The midpoint pathology

	Align(σ, L, Ft, L)	Σ	p	c	f
a. σσσ(σσ)	***, **, *; *	7	6	1	0
b. (σσ)σσσ	*; **, ***, ****	10	0	1	9
☞ c. σσ(σσ)σ	**, *; *; **	6	3	1	2
d. σ(σσ)σσ	*; *; **, ***	7	1	1	5

As Hyde (2012) correctly points out, the problem here is not the gradience of Alignment alone, but only in connection with its relation-generality. I.e., for assessment of violations it doesn't matter in which relation the two arguments, here syllable and foot, are, whether one is contained in the other or not. In tableau (20), one violation is incurred in all candidates by the second syllable within the foot. If one looks at the two other types of syllable, i.e., preceding and following the foot, one sees variation in the violation profiles.

If one has a look at a range of Alignment constraints proposed in the literature, a striking property they all share is that there is an implicit containment relation between arguments. For categories from the prosodic hierarchy it is usually intended that the constraint assesses violations only for those structures in which the lower category is contained within the higher one (e.g., feet contained within a PWd are aligned with an edge of this PWd and not a neighbouring one). Hyde (2012) redefines Alignment in a way that specifies the alignment categories, the separator categories and the (containment) relations between them, such that irrelevant categories (such as the neighbouring PWd) are excluded from the computation of violation marks. Since the intervening category is explicitly defined, the locus of each constraint violation can be identified as a different one, as considered a defining property of categorical violation by McCarthy (2003). Thus, even though the Relation-Specific Alignment constraints potentially still measure distance, they are not gradient – at least not in the same way as those constraints discussed above, i.e., that the same item causes registration of more or less violation.

While some properties of constraints have dramatic consequences for OT computation regardless of assumptions on representations, others are crucially dependent on the theory of representation one adopts. However, even without commitment to a certain theory, as we have seen here, a lot of fruitful discussion is possible on the more general properties of OT constraints, and the discussion is an ongoing one. We can expect new developments regarding the choices between different positional theories (Faithfulness, Markedness, Licensing) or their hybridization. A hybrid between Faithfulness and Alignment has also been proposed with Anchoring constraints (McCarthy & Prince 1995). The above revisions to Alignment will surely have repercussions for Anchoring, which is commonly invoked in analyses of prosodic morphology (truncation, reduplication etc.).

3.4 Constraint interaction I: on the relative strictness of domination

As already alluded to in the introductory section, the standard ranking relation between constraints is strict domination. The number of violations on lower ranked constraints doesn't matter if a higher ranked constraint has forced a decision on two competitors, be these lower violations all violations of the same, or several, constraints.

Instead of strict domination one could also consider candidates to be evaluated by overall score. I illustrate this with the constraint causing final devoicing, VOP/coda, its more general sister VOP (Voiced Obstruent Prohibition), and the conflicting one blocking it, i.e., IO-Ident(voice), since they will become relevant again in the discussion of Pater's (2016) argument for Harmonic Grammar (HG) below. Constraint violations are indicated as negative numbers rather than asterisks in the following tableaux for ease of interpretation.

(21) Evaluation by total score, no prioritization (i.e., unranked/unweighted constraints)

/bagdabgad/	IO-Ident(voice)	VOP/coda	VOP	Σ
a. bagdabgad		−3	−6	−9
b. bakdabgad	−1	−2	−5	−8
c. bakdapgad	−2	−1	−4	−7
☞ d. bakdapgat	−3		−3	−6
☞ e. bakdapkat	−4		−2	−6
☞ f. baktapkat	−5		−1	−6
☞ g. paktapkat	−6			−6
h. pagtabkad	−3	−3	−3	−9

Total scores alone are also too restrictive for typological theory since all candidates would score the same in all languages. We can combine the numerical values with ranking by giving constraints different weights, as proposed in HG. This results in potential typologies, as illustrated by the tableaux with different weightings of the same constraints.

(22) HG – Contrast with weighted constraints

	IO-Ident(vce)x3	VOP/Codax1	VOPx1	Σ
☞ /pad/ – pad		−1	−1	−2
/pad/ – pat	−1			−3
☞ /bad/ – bad		−1	−2	−3
/bad/ – bat	−1		−1	−4

(23) HG – Positional neutralization with weighted constraints

	IO-Ident(vce)x2	VOP/Codax3	VOPx1	Σ
/pad/ – pad		−1	−1	−4
☞ /pad/ – pat	−1			−2
/bad/ – bad		−1	−2	−5
☞ /bad/ – bat	−1		−1	−3

From a computational perspective, constraint weighting is more costly than strict domination, since it needs numerical calculation, while strict domination is digital in the sense that counting is needed only up to 1. (Either a candidate has one more violation of a constraint than a competitor or it doesn't.) Thus, one would expect a good argument for HG, which would be that it is either more restrictive than strict domination or it can account for patterns with which strict domination struggles. Pater (2016) claims gang-up effects to be the most convincing argument for weighted constraints. He provides an analysis of Lyman's Law effects in loanwords in Japanese (see Ito & Mester 2003; Kawahara 2011).

In recent Japanese loanwords, voiced geminates are allowed (while not attested in the native vocabulary) and only devoiced if they also violate Lyman's Law (no two voiced obstruents within a word). The latter is also only enforced in these loanwords when a voiced geminate is involved (in addition to a second voiced obstruent). Pater proposes a very elegant analysis of these data utilizing weighted constraints. At the core of this analysis lies the cumulative violation of the Lyman's Law constraint (OCP-Voice in (24)) and the M constraint against voiced geminates as more important than violation of an F constraint that has a higher weight than each of the two M constraints.

(24) HG analysis of Lyman's Law and OCP conspiracy (adapted from Pater 2016)

/dog:u/	Ident-Voice 3	OCP-Voice 2	*Vce-Gem 2	Σ
a. dog:u		−1	−1	−4
☞ b. dok:u	−1			−3
c. tog:u	−1		−1	−5
d. tok:u	−2			−6

However, he also shows at length that gang-up effects are an expected result in HG. Let us investigate briefly what a gang-up effect is. With strict domination, candidate (b) wouldn't have won against candidate (a), since it violates the higher ranked Ident-Voice. Here it is assumed that the two constraints with less weight, OCP-Voice and *Vce-Gem "gang up" against Ident-Voice. Their cumulative violations of candidate (a) outweigh the only relatively weighty violation of Ident-Voice incurred by candidate (b). In a gang-up, the violations of one or two less important constraints outweigh the violations of a higher ranked constraint. This kind of effect is also logically possible in other phonological phenomena. However, it is not attested. A gang-up could lead to inconsistency in final devoicing or other neutralization patterns. Also in assimilation patterns one could logically expect that assimilation results in the shared feature value that is held by the majority of involved segments in the input, i.e., the "majority rules" (Baković 2000). I first consider voicing neutralization and assimilation and then turn to vowel harmony.

If we, for the moment, stick to obstruent voicing patterns, which Pater (2016) also uses in his argumentation for weighted constraints, final devoicing in interaction with voicing assimilation could be analyzed as in the following tableaux. The usual pattern, as found in Dutch, Russian and many other languages, is regressive assimilation. Final devoicing is overridden by assimilation. In the following tableaux I will use the positional Faithfulness approach of final devoicing, since that easily accounts for the regressive nature of assimilation in connection with final devoicing (Lombardi 1999; see, e.g., Krämer 2000; Grijzenhout & Krämer 2000 for some discussion and Brown (2016) for an updated typology of voicing patterns).

(25) HG analysis of regressive voicing assimilation and final devoicing

i. /abga/	Agree 4	IdentOns 5	VOP 2	Ident 1	Σ
☞ abga			−2		−4
apga	−1		−1	−1	−7
apka		−1	0	−2	−7
ii. /apga/	Agree 4	IdentOns 5	VOP 2	Ident 1	Σ
☞ abga			−2	−1	−5
apga	−1		−1		−6
apka		−1	0	−1	−6
iii. /abka/	Agree 4	IdentOns 5	VOP 2	Ident 1	Σ
abga		−1	−2		−9
apga	−1	−1	−1	−1	−14
☞ apka				−1	−1

As the observant reader will have noticed, there is quite a safety distance in the weights of the top-weighted constraints. If we go for minimal weight differences the pattern turns out differently. In the following tableaux the weights are minimally different from each other and Agree weighs heavier than IdentOnset. The result is devoicing whenever at least one of the input segments is voiceless.

(26) Directionality switch dependent on underlying specification of one C

i. /abga/	Agree 4	IdentOns 3	VOP 2	Ident 1	Σ
☞ abga			−2		−4
apga	−1		−1	−1	−7
apka		−1		−2	−5
ii. /apga/	Agree 4	IdentOns 3	VOP 2	Ident 1	Σ
abga			−2	−1	−5
apga	−1		−1		−6
☞ apka		−1	0	−1	−4
iii. /abka/	Agree 4	IdentOns 3	VOP 2	Ident 1	Σ
abga		−1	−2		−7
apga	−1	−1	−1	−1	−10
☞ apka				−1	−1

If we reverse the relation between the two top-weighted constraints, the pattern generated by this grammar displays free variation in cases in which the input contains a combination of a voiceless and a voiced obstruent, in that order (see sub-tableau (27)ii).

(27) Free variation dependent on underlying specification of Cs

i. /abga/	Agree 3	IdentOns 4	VOP 2	Ident 1	Σ
☞ abga			−2		−4
apga	−1		−1	−1	−6
apka		−1		−2	−6
ii. /apga/	Agree 3	IdentOns 4	VOP 2	Ident 1	Σ
☞ abga			−2	−1	−5
☞ apga	−1		−1		−5
☞ apka		−1		−1	−5
iii. /abka/	Agree 3	IdentOns 4	VOP 2	Ident 1	Σ
abga		−1	−2		−8
apga	−1	−1	−1	−1	−10
☞ apka				−1	−1

Another problematic potential gang-up effect is that independent violations of a single constraint can add up to outweigh those of more weighty constraints in competing candidates. Legendre, Sorace & Smolensky (2006) refer to this as an unbounded trade-off. Here several items within a candidate "gang up" against another one rather than two or more constraints joining forces.

This is illustrated here with vowel harmony (VH; see Legendre, Sorace & Smolensky 2006 for an example involving stress placement). VH systems are often of the controlled type (as opposed to dominance of one feature value). In controlled systems a vowel in a prominent position – the stem, the first syllable, the stressed syllable or the rightmost syllable (see Baković 2000; Krämer 2003a) – causes all other vowels to assimilate. This is illustrated in (28) in sub-tableau (i). However, in such an HG analysis, if a form contains too many vowels, their cumulative unfaithfulness can reverse the pattern. The prominent position becomes unfaithful to avoid too much unfaithfulness among non-prominent vowels, as illustrated in tableau (ii) in (28).

(28) Majority rules in vowel harmony

i. /$\overset{+}{V}_p\overset{-}{V}\overset{-}{V}$/	Agree 10	Faith/P 3	Faith 1	Σ
$\overset{+}{V}_p\overset{-}{V}\overset{-}{V}$	−1			−10
☞ $\overset{+}{V}_p\overset{+}{V}\overset{+}{V}$			−2	−2
$\overset{-}{V}_p\overset{-}{V}\overset{-}{V}$		−1	−1	−4

ii. /$\overset{+}{V}_p\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}$/	Agree 10	Faith/P 3	Faith 1	Σ
$\overset{+}{V}_p\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}$	−1			−10
$\overset{+}{V}_p\overset{+}{V}\overset{+}{V}\overset{+}{V}\overset{+}{V}\overset{+}{V}$			−5	−5
☞ $\overset{-}{V}_p\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}\overset{-}{V}$		−1	−1	−4

HG predicts an infinite number of languages in which the majority takes over at different points, ranging from two to an infinity (minus one) of underprivileged vowels ganging up.

However easy these patterns are to model in HG, VH systems are either of the controlled (by the stem or stressed position) or the dominant type (Baković 2000). This “gang-up” or “majority rules” VH type is unattested.[2]

The conclusion thus has to be that either HG shows that certain language types are possible, but currently unattested by coincidence, or that HG needs further stipulations on the weighting of constraints, which makes the theory less attractive, given that cases like the Japanese conspiracy alluded to above can be analyzed with different means as well.

Pater (2016, as McCarthy 2007) argues against LCC (Local Constraint Conjunction; see section 3.5) by constructing weird LCCs and showing how the theory overgenerates and how weighted constraints don’t overgenerate in the same way. Weighted constraints overgenerate in other ways. And, as indicated already in section 3.3, other subtheories within OT lead to overgeneration. At the current stage the hard truth to face for phonologists is that the challenge lies in accounting for the attested rather than excluding the unattested. The OT tool of factorial typology (see Iosad, this volume), considered the litmus test for any proposed constraint set, requires the analyst to consider all possible rankings of a set of constraints. In the ideal case, the different rankings describe different patterns, and all predicted patterns are attested and each attested pattern is generated by at least one ranking. However, as proponents of substance-free phonology point out, the set of currently known languages and patterns is not necessarily the same as the set of possible grammars or the set of grammars that a phonological theory is expected to account for (Hale & Reiss 2000, 2008; Reiss, this volume).

We thus need different evaluation metrics for competing theories or apply those we have in a different way. For example, while LCC is a logical option within the formal apparatus of OT, weighted constraints are a completely different hypothesis about constraint interaction, i.e., an entirely different conceptualization of ranking, which comes with its own set of additional stipulations (such as exponential increase in constraint weight).

At the start of this section I showed tableau (21) with unweighted constraints in which several candidates tie. One can interpret this as a result, i.e., the top-scoring candidates are in free variation. There has been a considerable amount of research on phonological variation in the sense of optionality in recent years, which resulted in various revisions to the theory of constraint interaction. Some phonological processes apply only optionally, such as final *t* deletion in English or Gorgia Toscana, the spirantization of voiceless stops in postvocalic position in Tuscan Italian (see Iosad, this volume, and Ramsammy, this volume for more details). In the original version of OT, free variation is excluded, since every constraint ranking has to be exhaustive. For constraints for which a language doesn’t provide a ranking argument, a random or default ranking has to be assumed (e.g., M above F or specific above general). With a total ranking and every candidate supplied by Gen minimally differing from every other candidate by one violation mark, this results in one and only one winner in each evaluation.

There are several competing proposals to account for free variation in OT. The most straightforward approach is Partially Ordered Grammars Theory (Anttila 1997, 2004, 2007). Anttila still assumes exhaustive ranking. Though, this is enforced only temporarily. Constraints that are unranked with respect to each other assume a random order in every evaluation. Thus, considering two unranked constraints, *A* and *B*, the probability of constraint *A* dominating constraint *B* in an evaluation is 0.5. If each of the involved constraints favours a different of two candidates which otherwise tie, the chances of one or the other winning are 50%. Add a third constraint *C*, which favours the same candidate as constraint *A*, and the chances of this candidate to be chosen as optimal increase to 66%. In this way, free variation can be described, as well as frequency biases. The tableaux in (29) illustrate this schematically. Tableau (i) shows the unordered constraints and the subsequent tableaux the factorial typology that emerges with spontaneous rankings. In this scenario candidate *a* wins in four out of six possible rankings, *b* in two and *c* never. Thus, *a* and *b* are in free variation, with a higher likelihood of realization for *a*.

(29) Partially ordered constraints

i.		A	B	C
☞	a			*
☞	b	*	*	
	c	*	*	*

ii.	A	B	C
☞			*
	*!	*	
	*!	*	*

iii.	A	C	B
☞		*	
	*!		*
	*!	*	

iv.	B	A	C
☞			*
	*!	*	
	*!	*	*

v.		B	C	A
☞	a		*	
	b	*!		*
	c	*!	*	*

vi.	C	A	B
	*!		
☞		*	*
	*!	*	*

vii.	C	B	A
	*!		
☞		*	*
	*!	*	*

The same results can be achieved with Stochastic OT (Boersma 1998) and Maximum Entropy Grammar (Hayes & Wilson 2008) or Noisy Harmonic Grammar (Coetzee & Pater 2008), only with more complex maths involved in the computation.

A very different kind of softening up of the strictness of domination comes with the relation between constraints in content. At its most extreme, two constraints have the same content, which is only restricted to a subset of environments in one of them. This is the special-general relationship between constraints. This relation comes in various incarnations, which (probably) all can be summarized as constraint cloning. This will be discussed in slightly more detail in the next section, but see as well section 3.3 above.

3.5 Constraint interaction II: organization beyond ranking

In section 3.3, the problem with gradient or scalar constraints was introduced. The pointlessness of gradient violation under the strict ranking hypothesis did not only lead to proposals of different approaches to ranking but also to universal (strict dominance) rankings as well as to organizational relations between constraints beyond simple ranking, such as de Lacy's (2006) stringency relations. Gradience, however, is not the only problem that sparked the development of more sophisticated modes of constraint interaction and coordination or conjunction, and is not the actual problem, which are universal implicational (markedness) hierarchies. Other types of interaction are found in constraint duplication or cloning.

There are three ways of cloning constraints. Pater (2009) introduces the term in connection with lexical indexation of constraints. Constraints can be indexed, and the copy of the constraint with the index is ranked higher in a hierarchy than its unranked original. Some lexical items are also indexed, and it is only the output candidates of these inputs that are sensitive to the higher ranked indexed constraint clone. Positional Faithfulness can be understood as one type of cloning and indexing: There is a general F constraint and a more restricted version that is only active in a certain environment. There are two differences between the two forms of indexing. First, positional restriction is not arbitrary, as lexical indexing can be. It refers to well-defined positions or classes, such as stressed syllables (i.e., prosodically defined) or stems (i.e., morphologically defined). The boundary between the two types of indexing already becomes blurry in the latter case. For prosodically defined positions one could say that positional Faithfulness is defined over surface categories, while indexing is defined over input properties. A morphologically defined category, such as "stem", however, is clearly a property of the underlying form or input.[3]

The third cloning option is Local Constraint Conjunction (LCC; Prince & Smolensky 1993/2004; Smolensky 1996, 2006; Lubowicz 2002, 2005). Two (or more) constraints join forces in a domain (e.g., the segment) and every instance of that domain in which each of the two constraints is violated constitutes a violation of the local conjunction of the two (or more) constraints. The LCC only has an effect on output forms if it dominates at least one of the two constraints involved. However, indexation of an LCC with an arbitrary index or a general grammatical category (e.g., lexical vs. functional or major lexical class, as for positional Faithfulness or indexation) is an option.

3.5.1 Implicational hierarchies, gradience and constraint coordination

Freely ranked nuclear constraints seem to be badly suited to account for universal implicational relations or hierarchies. A famous exception is the relation between the constraints Onset and *Coda, which expresses a typological observation about syllable inventories (languages with syllables that have codas also have syllables with onsets, but not vice versa). This relation is captured in the respective positive and negative formulation of the constraints.

PoA in consonants shows a structurally comparable asymmetry. The three major PoAs are labial, coronal and dorsal. However, if a language has only two PoA, one usually finds a labial and a dorsal. Furthermore, many phonological processes indicate that coronal is the least marked of the three PoA, or even underspecified (see de Lacy 2006 and references given there), and glottal is even less marked than coronal, since it is the output of debuccalization and consonant epenthesis.

De Lacy (2006), based on previous work by Lombardi and many others, proposes the markedness hierarchy for PoA given in (30). These markedness relations between the different PoAs should be reflected in the constraint hierarchy, as in (b) or in the definition of the constraint set, as in (c). The constraints in (c) are in a stringency relation (de Lacy 2006). {*Dorsal/*Labial}, for example, is violated by any segment that is either dorsal or labial.[4]

(30) Markedness of Place of Articulation
  a. Dorsal > labial > coronal > glottal
  b. Universally ranked M constraints: *Dorsal ›› *Labial ›› *Coronal

c. Stringent constraint sets:
c'. *Dorsal, {*Dorsal/*Labial}, {*Dorsal/*Labial/*Coronal}
c''. Faith(dors), Faith(dors/lab), Faith(dors/lab/cor), Faith(d/l/c/glottal)

These stringently related constraints can be freely ranked and still express the markedness imbalance. Any dorsal in a candidate violates three M constraints, while a labial violates only two and so on.

The scalar behaviour discussed above relates to the sonority hierarchy. The problem there was that languages use different levels on the hierarchy as cut-off points for various processes. Prince & Smolensky (1993/2004) discuss syllable nuclei in this respect. While some languages allow only vowels, others allow sonorant consonants or even obstruents as syllable nuclei. However, every language always also allows the classes higher on the sonority hierarchy than the lowest one that is acceptable as a syllable nucleus in the respective language. Accordingly Prince & Smolensky decompose the gradient constraint H-Nuc into a universally fixed hierarchy of categorical M constraints. For a constraint on nuclear harmony one could assume the following gradient violation profile.

(31) Sonority and cumulative violation of H-Nuc

vowel	liquid	nasal	fricative	stop
√	*	**	***	****

However, no amount of violations of H-Nuc will ever trigger vowel epenthesis to provide a better nucleus for a syllable as long as the anti-epenthesis constraint Dep is ranked above H-Nuc.

(32) H-Nuc decomposed
*stop/nuc ›› *Fric/nuc ›› *Nas/nuc ›› *Liq/nuc

Dep can now be ranked somewhere in between these constraints and block vowel epenthesis in forms in which any of the classes referred to lower down in the hierarchy is syllabified as a nucleus. The universal ranking of these constraints undermines the Free Ranking Hypothesis, and it thus would be desirable to reformulate the same insight with freely rankable categorical constraints in the same way as the constraints on the major PoAs.

(33) H-Nuc stringently decomposed
*stop/nuc
*{Stop∨Fric}/nuc
*{Stop∨Fric∨Nas}/nuc
*{Stop∨Fric∨Nas∨Liq}/nuc

Similar issues are identifiable with Faithfulness in chain shifts. In many lenition processes consonants only move up one step on the sonority hierarchy. Likewise, vowel raising often moves the vowel up only one step in the height dimension, e.g., low vowels raise to lax mid, lax mid raise to tense mid and tense mid raise to high, as in metaphony in Romance languages (see, e.g., Gnanadesikan 1997 and Moreton & Smolensky 2002 for an overview of chain shifts; Iosad 2010 on mutations; Calabrese 2011 or Krämer 2016 on metaphonic raising). One solution that has been proposed relies crucially on LCC (e.g., Kirchner 1996; Moreton & Smolensky 2002).

In an LCC two (or more) constraints team up and can be ranked higher than the individual constraints themselves.

(34) Local Constraint Conjunction (Smolensky 2006)
$*A\&_D*B$ is violated if and only if a violation of *A and a (distinct) violation of *B both occur within a single domain of type *D*.

Cooccurrence constraints, such as the constraint responsible for the observation that high lax vowels are typologically marked, are pretty complex constraints if conceived as primitive constraints. Decomposing them as an LCC of two primitive constraints yields a more elegant constraint set. I.e., high vowels violate the simple constraint *[+high] and lax vowels violate the constraint *[Retracted Tongue Root] (or *[−Advanced Tongue Root]). While high vowels are typologically very common and vowels with retracted tongue root are very common as well, the combination of both in one segment is marked. This is captured in the LCC $\{*[\text{high}]\&_{segment}*[\text{RTR}]\}$.

Kirchner (1996) proposes to handle chain shifts with LCCs of F constraints (see as well Krämer 2016; Walker 2011). A chain shift grammar tolerates violation of one F constraint, i.e., a change of one feature in one segment, but not two (or more) violations, i.e., two feature changes in the same segment. An issue that arises here is the formalization of the triggering M constraint, since often the goal of markedness reduction is not accomplished, it is only approached by one step.

LCC has repeatedly been criticized for being too powerful, allowing all kinds of undesired constraint interactions (see Pater 2016 for the latest assault), especially if one allows all sorts of domains beyond that of the segment, despite the proven usefulness and explanatory adequacy of LCCs in the analysis of a wide range of phonological phenomena (see, e.g., Kirchner 1996; Moreton & Smolensky 2002; Smolensky 2006; Collins 2013).

Furthermore, as pointed out already by Crowhurst & Hewitt (1997), conjunction is only one logical operation available for the coordination of constraints. De Lacy's (2006) stringently organized constraints, for example, can be analyzed as local disjunctions, i.e., a single segment should not violate either constraint *A* or constraint *B*, e.g., either *Labial *or* *Dorsal. Implicational constraints of the type "if x then also y" have also been proposed in various forms by Krämer (2003a: 86); Smith (2005); and Levelt & van Oostendorp (2007).

3.5.2 Exceptionality and constraint indexation

Phonological processes often only apply to restricted lexical classes or to arbitrarily selected individual lexical items or morphemes. In many languages, loanwords also form a separate phonological class in which different, often more loose, restrictions hold than in the native vocabulary (though see Jurgec 2010 for the contrary). The phenomenon of exceptionality and loanword phonology have first been handled with co-phonologies, i.e., the duplication of the complete constraint hierarchy, or the duplication and reranking of a substring of the hierarchy (Anttila 2002) or with prespecification (exceptional processes only apply to arbitrarily underspecified morphemes; Inkelas, Orgun & Zoll 1997 et seq.).

A more restrictive and more insightful theory of exceptions is constraint indexation (Itô & Mester 1999). A constraint is cloned, tagged with an index and placed higher up in the hierarchy than the original constraint. The indexed constraint is only visible, or only registers violation marks, for morphemes that are tagged with the same index. These morphemes can be loanwords, lexical categories or random groups of morphemes or single morphemes. If both M and F constraints can be indexed, as in Pater's (2009) version, the approach produces three welcome results. Not only does it distinguish loanwords from the native vocabulary (by associating certain F constraints and all loanwords with an index), it accounts for the observation that exceptionality is morphophonologically local (though see Jurgec 2014), and it distinguishes between exceptional blocking and exceptional application of a phonological process.

If two morphemes from different classes combine, it is not clear under the co-phonology approach which constraint ranking should be used. Furthermore, the presence of one morpheme from a co-phonology P' causes the whole form to be subdued to the constraint ranking of P'. Lexically indexed constraints are activated by the corresponding lexically indexed morphemes and only apply to these morphemes. That is, different co-phonologies can be activated within one morphologically complex form.

Whether the activity concerns the exceptional blocking or the exceptional application of a process depends on whether the indexed constraint is an F constraint or an M constraint, respectively.

(35) Indexation and exceptionality
a. Exceptional process: M_I ›› F ›› M
b. Exception to process: F_I ›› M ›› F
c. Loanword exceptionality: F_L ›› M ›› F

Lexical indices are of course simple diacritics, and therefore this is not a phonological solution to the challenge. Morpheme-specific processes and blocking analyses relying on under-/pre-specification or floating features seem to be more attractive since they don't have to rely on arbitrary diacritics. If, however, loanword phonology and lexical class-specific phonology also require indexation it is tempting to apply a uniform analysis to all forms of exceptionality.

3.6 Related topics and future directions

Unfortunately this chapter comes with a severe flaw: restricted space. While we covered some ground, this overview of current issues and new developments is far from exhaustive, and I use the final section to draw attention to some additional issues and trends.

Overgeneration has been touched in passing, even though it would have deserved its own section (see Iosad, this volume, for more discussion). Constraint interaction has been shown repeatedly to show undesired results. For example, Steriade (2009) notes that nasalization is not an attested repair strategy in response to a constraint like VOP/coda (the Final Devoicing constraint). However, a mapping of /tab/ to [tam] is easily produced with the respective ranking of common OT constraints. The too-many-solutions or too-many-repairs problem has been discussed in various places with very divergent results (Blumenfeld 2006; de Lacy 2006; Baković 2007a; van Oostendorp 2007b; Blaho & Rice 2014).

Overgeneration, in some sense, is tightly connected with computability, which has barely been mentioned so far. The infinite set of output candidates and their parallel evaluation poses a potential challenge which has been met with some scholars' move to a serialist version of OT or HG (though see Iosad, this volume). In serial OT only one change can be made at a time to the input, and evaluation is repeated until there are no minimal changes left that would improve the output with respect to the constraint ranking. On the one hand this results in a potentially long chain of evaluations; on the other it restricts the set of output candidates in each evaluation round dramatically. However, Magri (2013) raises serious doubts about the computational advantage of HG over OT. Bane & Riggle (2012) point out that HG generates much larger factorial typologies than standard OT. Kaplan (2011) and Kazutaka (2012) argue that Harmonic Serialism is problematic because it can't (straightforwardly) account for patterns that are elegantly taken care of with parallel OT. See as well Hyde (2012) for a critical stance on serial OT.

The adoption of a serial version of OT has its motivation in OT's problem with opaque interactions of phonological processes, which has sparked a firework of theoretical proposals. McCarthy (2007) gives an overview of the state of the discussion at that time. The issue is far from resolved, as subsequent contributions show, for example Padgett's (2010) reductionist approach, Baković's (2007b, 2011) reassessment of the phenomenon or van Oostendorp's (2004, 2017) return to the original Containment model of Faithfulness with slight modifications, Coloured Containment. See as well Bermúdez-Otero (this volume).

The functional move that came with OT has led to the inclusion of phonetic detail, for example in the definition of constraints, such as formant frequencies, and the ambition to explain a much greater amount of variation than discussed above, i.e., phonetic variation. This raises the question of whether OT is actually a theory of competence or performance (in the Chomskyan sense) or both.

The inclusion of phonetic detail (rather than abstract categorical features) in phonological constraints is also relevant for another discussion that is orthogonal to OT-internal discussions: the nature of underlying representations. While traditional generative phonology endorses abstract categorical phonological representations, there has been an ongoing discussion, especially since the nineties of the last century, concerning the degree of abstractness and economy in underlying or lexical representations, i.e., the mental representation of phonological objects, or, in a wider view, speech (see summarizing discussion and references in Krämer 2012). The debate (within OT) was sparked already in Prince & Smolensky (1993/2004) by their discussion of Lexicon Optimization and its undesired results and has led to a range of proposals (for more recent contributions, see Krämer 2012; Tesar 2013; van Oostendorp 2014).

In conclusion, it seems there are not many of the basic assumptions of OT that are not under debate, and it is going to be interesting to see how the framework will develop in the future.

Further reading

While learnability was a central issue from the beginning on, and considered one of the strong arguments in favour of OT, it is still a hotly debated issue, spawning new proposals, e.g., Brasoveanu & Prince (2011); Tesar (2013); Tessier & Jesney (2014); and Rasin & Katzir (2016).

The most recent trend in OT has turned to a more fine-grained investigation of the typological properties of systems of rankable constraints, e.g., Alber, Busso & Prince (2016); Brasoveanu & Prince (2011); or McManus (2016) and the assumed candidate sets, e.g., Bane & Riggle (2012). With these projects research in OT has explicitly turned its focus from phonological phenomena, such as opacity or gradience, to the theory itself and its properties, as the subject of investigation.

Notes

1 The choice of fricative surfacing for the /g/ is subject to the *ich*-*ach*-Laut alternation, which is irrelevant here.
2 Admittedly, such gang-up effects caused by F constraints are excluded in a serial version of HG by the restriction to one change (i.e., one violation of one F constraint) in the propagation from one

representation to the next in an evaluation. However, for iterative assimilation patterns this theory then requires an analysis crucially relying on gradient Alignment (see the discussion here and in McCarthy 2003 and Hyde 2012 on why this is a problem and Jurgec 2011 why it isn't) as the driving force for vowel harmony and other unbounded assimilation processes, since, for AGREE constraints, a single change that doesn't necessarily increase harmony, e.g., the sequence of [++ − −], is as disharmonic as the sequence [+ − − −] or [+++ −]. The grammar thus wouldn't be able to select an appropriate input for the next evaluation.

3 Especially so if one embraces strict modularity and all syntactic, semantic and morphological – that is, non-phonological – information is considered inaccessible in the phonological computation.

4 Iroquoian languages that don't have any labial consonants in their inventory, such as Seneca (Chafe 1996), are potentially difficult for Dispersion Theory (Flemming 2004), though easily accounted for with de Lacy's constraints on PoA.

References

Alber, Birgit, Natalie Del Busso & Alan Prince 2016. From intensional properties to universal support. *Language* 29(2): e88–e116. DOI: 10.1353/lan.2016.0029.

Anttila, Arto 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & W. Leo Wetzels (eds.) *Variation, Change and Phonological Theory*. Amsterdam/Philadelphia: Benjamins, 35–68.

Anttila, Arto 2002. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory* 20: 1–42.

Anttila, Arto 2004. Variation and phonological theory. In J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *The Handbook of Language Variation and Change*. Oxford: Blackwell Publishing, 206–243.

Anttila, Arto 2007. Variation and optionality. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 519–536.

Archangeli, Diana & Douglas Pulleyblank 2015. Phonology without universal grammar. *Frontiers in Psychology* 6: article 1229.

Baković, Eric 2000. *Harmony, dominance and control*. PhD thesis, UMass. ROA 360.

Baković, Eric 2007a. Local assimilation and constraint interaction. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 335–352.

Baković, Eric 2007b. A revised typology of opaque generalisations. *Phonology* 24: 217–259.

Baković, Eric 2011. Opacity and ordering. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (eds.) *The Handbook of Phonological Theory*, 2nd ed. Oxford: Blackwell, 40–67.

Bane, Max & Jason Riggle 2012. Consequences of candidate omission. *Linguistic Inquiry* 43: 695–706.

Beckman, Jill 1997. Positional faithfulness, positional neutralisation, and shona vowel harmony. *Phonology* 14(1): 1–46.

Beckman, Jill 2004. The case for local conjunction: Evidence from Fyem. In Mimu Tsujimura & Gina Garding (eds.) *WCCFL 22 Proceedings*. Somerville, MA: Cascadilla Press, 56–69.

Bermúdez-Otero, Ricardo 2011. Cyclicity. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell Companion to Phonology*, vol. 4. London: Wiley-Blackwell, 2019–2048.

Bermúdez-Otero, Ricardo forthcoming. *Stratal Optimality Theory*. Oxford: Oxford University Press.

Blaho, Sylvia & Curt Rice 2014. Overgeneration and falsifiability in phonological theory. In Jacques Durand, Gjert Kristoffersen, Bernard Laks (eds.) *avec la collaboration de Julie Peuvergne. La phonologie du français: normes, périphéries, modélisation. Mélanges pour Chantal Lyche*. Presses Universitaires de Paris Ouest, 101–120.

Blevins, Juliette 1995. The syllable in phonological theory. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell, 206–244.

Blumenfeld, Lev 2006. *Constraints on phonological interactions*. PhD dissertation, Stanford.

Boersma, Paul 1998. *Functional phonology – Formalizing the interactions between articulatory and perceptual drives*. PhD thesis, University of Amsterdam, Amsterdam.

Brasoveanu, Adrian & Alan Prince 2011. Ranking and necessity. *Natural Language & Linguistic Theory* 29: 3–70.
Broselow, Ellen, Su-I Chen & Chilin Wang 1998. The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20: 261–280.
Brown, Jason 2016. Laryngeal assimilation, markedness and typology. *Phonology* 33: 393–423.
Calabrese, Andrea 2011. Metaphony in Romance. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell Companion to Phonology*. London: Blackwell, 2631–2661.
Chafe, Wallace L. 1996. Sketch of Seneca, an Iroquoian language. In William C. Sturtevant (ed.) *Handbook of North-American Indians. Volume 17: Languages*, ed. by Ives Goddard. Washington: Smithsonian Institution, 551–579.
Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
Coetzee, Andries & Joe Pater 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26: 289–337.
Collins, Joe 2013. *Modal-dependence and naturalness in phonology: Confronting the ontogenetic question*. Master's thesis, UiT the Arctic University of Norway.
Crowhurst, Megan & Marc Hewitt. 1997. Boolean operations and constraint interactions in Optimality Theory. Ms., University of North Carolina at Chapel Hill and Brandeis University.
de Lacy, Paul 2006. *Markedness*. Cambridge: Cambridge University Press.
Flemming, Edward 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.) *Phonetically-Based Phonology*. Cambridge: Cambridge University Press, 232–276.
Fortescue, Michael 1984. *West Greenlandic*. London: Croom Helm.
Gnanadesikan, Amalia 1997. *Phonology with ternary scales*. Doctoral dissertation, University of Massachusetts, Amherst.
Goldsmith, John 1990. *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.
Goldsmith, John 1993. Harmonic phonology. In John Goldsmith (ed.) *The Last Phonological Rule*. Chicago: University of Chicago Press, 21–60.
Gordon, Matthew 2007. Functionalism in phonology. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 61–78.
Grijzenhout, Janet & Martin Krämer 2000. Final devoicing and voicing assimilation in Dutch derivation and cliticization. In Barbara Stiebels & Dieter Wunderlich (eds.) *Lexicon in Focus*. Studia Grammatica 45. Berlin: Akademie Verlag, 55–82.
Hale, Mark & Charles Reiss 2000. "Substance abuse" and "dysfunctionalism": Current trends in phonology. *Linguistic Inquiry* 31: 157–169.
Hale, Mark & Charles Reiss 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
Hayes, Bruce 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In Michael Darnell, Edith Moravcsik, Frederick J. Newmeyer, Michael Noonan, & Kathleen Wheatley (eds.) *Functionalism and Formalism in Linguistics*. Amsterdam and Philadelphia: John Benjamins, 243–285.
Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.) 2004/2008. *Phonetically Based Phonology*. Cambridge: Cambridge University Press.
Hayes, Bruce & Colin Wilson 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440.
Hyde, Brett. 2012. Alignment constraints. *Natural Language and Linguistic Theory* 30: 789–836.
Inkelas, Sharon, Orhan Orgun & Cheryl Zoll 1997. The implications of lexical exceptions for the nature of grammar. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon Press, 393–418.
Iosad, Pavel 2010. Right at the left edge: Initial consonant mutations in the languages of the world. In Michael Cysouw & Jan Wohlgemuth (eds.) *Rethinking Universals: How Rarities Affect Linguistic Theory*. Berlin: Mouton de Gruyter, 105–138.
Itô, Junko 1988. *Syllable Theory in Prosodic Phonology*. New York: Garland Publishing.

Itô, Junko & Armin Mester 1999. The structure of the phonological lexicon. In Natsuko Tsujimura (ed.) *The Handbook of Japanese Linguistics*. Oxford: Blackwell, 62–100.
Jurgec, Peter 2010. Disjunctive lexical stratification. *Linguistic Inquiry* 41: 149–161.
Jurgec, Peter 2011. *Feature spreading 2.0. A unified theory of assimilation*. PhD thesis, University of Tromsø. Lingbuzz 001281.
Jurgec, Peter 2014. Morphology affects loanword phonology. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.) *Proceedings of NELS 43*, vol. 1. Amherst, MA: GLSA, 191–202.
Kaplan, Aaron 2011. Harmonic improvement without candidate chains in Chamorro. *Linguistic Inquiry* 42: 631–650.
Kawahara, Shigeto 2011. Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29: 705–723.
Kazutaka, Kurisu 2012. Fell-swoop onset deletion. *Linguistic Inquiry* 43: 309–321.
Kiparsky, Paul 1973. Abstractness, opacity, and global rules: Part 2 of phonological representations. In Osamu Fujimura (ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC Corporation, 57–86.
Kiparsky, Paul 2000. Opacity and cyclicity. *The Linguistic Review* 17: 351–367.
Kirchner, Robert 1996. Synchronic chain shifts in optimality theory. *Linguistic Inquiry* 27: 341–350.
Kochetov, Alexei & Connie K. So 2007. Place assimilation and phonetic grounding – a cross-linguistic perceptual study. *Phonology* 24: 397–432.
Krämer, Martin 2000. Voicing alternations and underlying representations: The case of Breton. *Lingua* 110(9): 639–663.
Krämer, Martin 2003a. *Vowel Harmony and Correspondence Theory*. Mouton: Walter de Gruyter.
Krämer, Martin 2003b. *What is wrong with the right side? Edge (a)symmetries in phonology and morphology*. Ms., University of Ulster. ROA 576.
Krämer, Martin 2009. *The Phonology of Italian*. Oxford: Oxford University Press.
Krämer, Martin 2012. *Underlying Representations*. Cambridge: Cambridge University Press.
Krämer, Martin 2016. Metaphonic chain shifts, vowel height, and markedness. In Francesc Torres-Tamarit, Kathrin Linke & Marc van Oostendorp (eds.) *Approaches to Metaphony in the Languages of Italy*. Berlin: Mouton, 277–299.
Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky 1990. *Can connectionism contribute to syntax?: Harmonic grammar, with an application*. Report CU-CS-485–90. Computer Science Department, University of Colorado at Boulder.
Legendre, Géraldine, Antonella Sorace & Paul Smolensky 2006. The Optimality Theory-Harmonic Grammar connection. In Paul Smolensky & Géraldine Legendre (eds.) *The Harmonic Mind: From Neural Computation to Optimality Theoretic Grammar, vol. 2: Linguistic and Philosophical Implications*. Cambridge, MA: MIT Press, 339–402.
Levelt, Claartje & Marc van Oostendorp 2007. Feature co-occurrence constraints in L1 acquisition. *Linguistics in the Netherlands* 24: 162–172.
Lombardi, Linda 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17: 267–302.
Lubowicz, Anna 2002. Derived environment effects in Optimality Theory. *Lingua* 112: 243–280.
Lubowicz, Anna 2005. Locality of conjunction. In John Alderete, Chung-hye Han & Alexei Kochetov (eds.) *Proceedings of the 24th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, 254–262.
Magri, Giorgio 2013. HG has no computational advantages over OT: Towards a new toolkit for computational OT. *Linguistic Inquiry* 44: 569–609.
Magri, Giorgio 2016. Error-driven learning in OT and HG: A comparison. *Phonology*, 493–532.
McCarthy, John J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 38: 169–195.
McCarthy, John J. 1996. Remarks on phonological opacity in Optimality Theory. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.) *Studies in Afroasiatic Grammar: Papers from*

the Second Conference on Afroasiatic Linguistics, Sophia Antipolis, *1994*. The Hague: Holland Academic Graphics, 215–243.
McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16: 331–399.
McCarthy, John J. 2000. Harmonic serialism and parallelism. In Masako Hirotani, Andries Coetzee, Nancy Hall & Ji-yung Kim (eds.) *Proceedings of the North East Linguistics Society*. Amherst, MA: GLSA, 501–524.
McCarthy, John J. 2003. OT constraints are categorical. *Phonology* 20: 75–138.
McCarthy, John J. 2005. Taking a free ride in morphophonemic learning. *Catalan Journal of Linguistics* 4: 19–56.
McCarthy, John J. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.
McCarthy, John J. & Joe Pater (eds.) 2016. *Harmonic Grammar and Harmonic Serialism*. London: Equinox.
McCarthy, John J. & Alan S. Prince 1993. Generalized alignment. In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 79–193.
McCarthy, John & Alan S. Prince 1994. The Emergence of the Unmarked. In *Proceedings of the Northeast Linguistic Society* 24, GLSA, University of Massachusetts, Amherst. 333–379.
McCarthy, John & Alan S. Prince 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *UMOP 18: Papers in Optimality Theory*. University of Massachusetts, Amherst: GLSA, 249–384.
McCarthy, John & Alan S. Prince 1999. Faithfulness and identity in prosodic morphology. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press, 218–309.
McManus, Hope Eliza 2016. *Stress parallels in modern OT*. PhD dissertation, Rutgers University.
Moreton, Elliott & Paul Smolensky 2002. Typological Consequences of Local Constraint Conjunction. In *Proceedings of WCCFL* 21. 306–319.
Oostendorp, Marc van 2004. *The theory of faithfulness*. Ms., Meertens Institute, Amsterdam.
Oostendorp, Marc van 2007a. Derived environment effects and consistency of exponence. In Sylvia Blaho, Patrik Bye & Martin Krämer (eds.) *Freedom of Analysis?* Berlin: Walter de Gruyter, 123–148.
Oostendorp, Marc van 2007b. *Restricting repairs*. Unpublished ms., Amsterdam: Meertens Institute.
Oostendorp, Marc van 2014. Selective Lexicon optimization. *Lingua* 142: 76–84.
Oostendorp, Marc van 2017. *Faithfulness in Phonological Theory*. London: Equinox.
Orgun, Orhan & Ronald Sprouse 2010. Hard constraints in Optimality Theory. In Curt Rice & Sylvia Blaho (eds.) *Modelling Ungrammaticality in Optimality Theory*. London: Equinox, 97–114.
Padgett, Jaye 2010. Russian consonant-vowel interactions and derivational opacity. In W. Brown, A. Cooper, A. Fisher, E. Kesici, N. Predolac & D. Zec (eds.) *Proceedings of the Eighteenth Formal Approaches to Slavic Linguistics meeting (Second Cornell Meeting*, *2009)*. Ann Arbor: Michigan Slavic Publications, 353–382.
Pater, Joe 2009. Morpheme-specific phonology as constraint indexation and inconsistency resolution. In Steve Parker (ed.) *Phonological Argumentation*. London: Equinox Publications, 123–154.
Pater, Joe 2016. Universal grammar with weighted constraints. In John J. McCarthy & Joe Pater (eds.) *Harmonic Grammar and Harmonic Serialism*. London: Equinox, 1–46.
Piggott, Glyne L. 1999. At the right edge of words. *The Linguistic Review* 16: 143–185.
Prince, Alan & Paul Smolensky 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. London: Blackwell.
Prince, Alan & Bruce Tesar 2004. Learning phonotactic distributions. In René Kager et al. (eds.) *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 245–291.
Rasin, Ezer & Roni Katzir 2016. On evaluation metrics in Optimality Theory. *Linguistic Inquiry* 47: 235–282.
Rischel, Jørgen 1974. *Topics in West Greenlandic Phonology*. Copenhagen: Akademisk forlag.

Rubach, Jerzy 1997. Extrasyllabic consonants in Polish: Derivational Optimality Theory. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Oxford University Press, 551–582.

Smith, Jennifer L. 2005. *Phonological Augmentation in Prominent Positions*. New York: Routledge.

Smolensky, Paul 2006. Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In Paul Smolensky & Géraldine Legendre (eds.) *The Harmonic Mind: From Neural Computation to Optimality Theoretic Grammar, vol. 2: Linguistic and Philosophical Implications*. Cambridge, MA: MIT Press, 27–160.

Steriade, Donca 2009. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In K. Hanson & S. Inkelas (eds.) *The Nature of the Word: Studies in Honor of Paul Kiparsky*. Cambridge, MA: MIT Press, 151–179.

Tesar, Bruce 2013. *Output-Driven Phonology: Theory and Learning*. Cambridge: Cambridge University Press.

Tessier, Anne-Michelle & Karen Jesney 2014. Learning in Harmonic Serialism and the necessity of a richer base. *Phonology* 31: 155–178.

Torres-Tamarit, Francesc, Kathrin Linke & Marc van Oostendorp (eds.) 2016. *Approaches to Metaphony in the Languages of Italy*. Phonology and Phonetics 20. Berlin: Mouton de Gruyter.

Walker, Rachel 2005. Weak triggers in vowel harmony. *Natural Language and Linguistic Theory* 23: 917–89.

Walker, Rachel 2011. *Vowel Patterns in Language*. Cambridge: Cambridge University Press.

Weijer, Jeroen van de 2012. *Combining Optimality Theory and Exemplar Theory*. Nagoya Japan: Kougaku.

Zoll, Cheryl 2004. Positional asymmetries and licensing. In John J. McCarthy (ed.) *Optimality Theory in Phonology: A Reader*. Oxford: Blackwell, 365–378.

4

The phonology–phonetics interface in constraint-based grammar

Michael Ramsammy

4.1 Introduction

4.1.1 Overview of the chapter

Picking up themes from the previous two chapters, the overarching aim of this chapter is to present a discussion of how constraint-based theories of grammar conceive of the interfaces shared by phonology and other modules of grammar. What are sometimes referred to as *interface phenomena* present particular challenges for phonological theory. Consequently, the need to account for sound patterns that arise because of the sharing of structure between phonology and other components of grammar has led to a great deal of theoretical innovation.

For example, in the previous chapter in this volume, Krämer (this volume) comments on data from varieties of German that exhibit /g/-spirantisation. The alternation between the stop, [g], and the fricatives, [x] and [ç], in words like [flux]~[fly.gə] 'flight.SG/PL' and [ʁe.gən]~[ʁeç.nən] 'rain.N/INF' is dependent upon syllable structure (i.e. a phonological factor): observe that the stop [g] occurs in ONSET position, whereas the fricative allophones occur in coda position. However, these differences in syllable structure are conditioned by morphological structure (i.e. affixation).

Accounting for morphologically driven phonological alternations of this sort in the languages of the world has been perhaps the most central concern of phonological theory throughout the 20th and 21st centuries; and the emergence of Optimality Theory (OT) in the 1990s opened up new possibilities for analysing morphologically conditioned alternations. Furthermore, there have been various changes proposed to the 'classic' OT model in order to equip constraint-based phonology with mechanisms for dealing with opaque phonological process that arise because of the interleaving of morphological and phonological structure (e.g. sympathy theory, OO-correspondence, Stratal OT, OT with candidate chains etc.: see Bermúdez-Otero, this volume; Krämer, this volume). Whilst all of these models depart from the original model proposed in Prince & Smolensky (1993) to some degree, they all retain the defining feature of OT: namely, that the phonological grammar is governed by constraint interaction.

But how far should this extend? Should OT be capable of accounting for all possible types of sound patterns that are demonstrably under the cognitive control of human speakers through the medium of constraint interaction? In addition to categorical alternations such as German /g/-spirantisation, should this also include gradient patterns? And what about variable patterns? Should constraint-based frameworks have the capacity to model speech-perception processes as well as speech-production processes?

These are not simple questions to answer. The issues of gradience vs categoricity and speech production vs perception are crucial, not only for discussions of synchronic phonology, but also for theories of sound change. Having a theory of how phonology and phonetics as modules of grammar communicate and interact synchronically and diachronically is, arguably, equally as important as having a theory of how morphological and phonological structures interact. In view of this, it should perhaps surprise us that considerably less attention has been paid to developing explicit theories of phonology–phonetics interactions in constraint-based grammar than to models capable of handling morphology–phonology interactions.

Nevertheless, this is not to say that the phonology–phonetics interface has been entirely ignored in OT: there are constraint-based theories that make serious and successful attempts at addressing the above questions. It is these frameworks that are the focus of this chapter. We shall examine the fundamental claims of two particular constraint-based models of phonology that have developed from rigorous theoretical enquiry into the nature of the relationship between phonology and phonetics. In section 4.2, I provide a brief overview of Stratal OT. This framework is well suited to accounting for interface phenomena for a number of reasons. First and foremost, the Stratal OT model proposed by Bermúdez-Otero (2007, 2011, this volume) is based on very specific assumptions about the architecture of grammar. These assumptions allow for explicit predictions to be formulated about how modules of grammar may interact. Secondly (and relatedly), this framework provides tools for analysing phonological phenomena that may be either categorical (e.g. cases of discrete allophony) or gradient (e.g. controlled coarticulation, partial lenition etc.). Thirdly, Stratal OT allows for explicit hypotheses to be formulated about how phonological processes emerge and mature over time. The *life cycle of phonological processes* capitalises upon the architectural assumptions needed to account for synchronic phonological alternations in Stratal OT for the purpose of explaining phonological change.

Having taken a tour of Stratal OT, section 4.3 turns to looking at stop spirantisation in greater detail. Here, rather than the German dialectal patterns, our case study is Tuscan Italian. A close look *Gorgia Toscana* (GT, hereafter) reveals a range of challenges for phonological analysis. For example, how can we account for the fact that spirantisation patterns vary across Tuscan dialects? And how is stylistic variation to be analysed? What can we say about the relationship between synchronic patterns of variation and historical change? After reflecting on these issues, sections 4.4–4.6 present a reanalysis of GT. Drawing upon the theoretical claims of the life cycle of phonological processes and of *Bidirectional Phonology and Phonetics* (Boersma 2007, 2009; Boersma & Hamann 2008; Hamann 2010), this analysis locates Tuscan spirantisation at the heart of the phonology–phonetics interface. In essence, this approach relates the patterns of variation observed in Tuscan spirantisation to both synchronic and diachronic interactions between phonological and phonetic modules of grammar. Section 4.7 concludes the chapter.

4.1.2 Beyond constraint-based phonology?

It is important to highlight at the outset that this chapter focuses narrowly on constraint-based models of grammar. This means that the analyses presented here are influenced by OT thinking in very theory-specific ways. Thus, the analyses discussed in this chapter exploit *one* conception of the phonology–phonetics interface: there are, of course, other ways of thinking about how phonology and phonetics interact, and what the division of labour between phonology and phonetics should be. As Hamann (2010: 222–223) notes, there can probably be no theory-neutral way of approaching a discussion of the phonology–phonetics interface. This chapter can therefore be most usefully read alongside other work that makes fundamentally different assumptions about how phonological and phonetic structures interact.

For relevant discussion, see Scobbie (2007) and chapters in this volume on Exemplar Theory, Phonetically Grounded Phonology and Articulatory Phonology.

4.2 Grammatical architecture and Stratal OT

4.2.1 Phonology as a component of grammar

In aiming to provide an overview of how phonological structures interact with other grammatical structures – i.e. in order to describe phonology and its interfaces – we must first define how and where phonology fits in to the larger grammatical picture. OT, both in its mainstream parallel version and the stratified version discussed below, is a theory of phonological computation that can be grouped with other theoretical frameworks that assume a modular grammatical architecture (Pierrehumbert 2002: section 4.2; Bermúdez-Otero 2012). This concept is schematised in Figure 4.1 below. In speech production, phonological structures stored in long-term memory are first retrieved from the lexicon and then 'fed forward' to the grammar. Like its rule-based predecessor, Lexical Phonology and Morphology (LPM hereafter: Kiparsky 1982, 1985; Mohanan 1986), Stratal OT assumes a close affinity between phonological and morphological structure. For the purposes of this chapter, we shall work from the assumption that it is in the morpho-phonological module that morphological operations (i.e. inflection and derivation) and all categorical phonological processes (i.e. operations involving discrete representational changes) take place. Outputs generated by the cumulative application of morpho-phonological processes are then fed forward again to the phonetic implementational module. In this module, discrete, feature-based phonological representations are converted into non-discrete phonetic objects (i.e. auditory representations and gestural implementation plans: see Keating 1990; Cohn 1993).

The process by which discrete phonological structures are converted into phonetic objects is termed *transduction* (or *translation*, alternatively). Transduction is the focal point of phonology–phonetics interactions in any theory assuming a modular structure to the grammar. Thus, much of this chapter is devoted to discussing how transduction processes can be conceived of in a constraint-based framework. In anticipation of this discussion, note that an important property of the interfaces between grammatical modules in this type of model is *bidirectionality* (represented by the dual-headed arrows in Figure 4.1). It is worth pointing out here that interfaces play a central role not only in speech production (i.e. feed-forward, left-to-right computation), but also in speech comprehension (i.e. 'feed-back' or right-to-left computation) and language acquisition. Some aspects of bidirectional interactions between the phonological and phonetic modules are discussed in sections 4.5–4.6; however, a detailed discussion of the role of such interactions in phonological learning is beyond the scope of this chapter.[1]

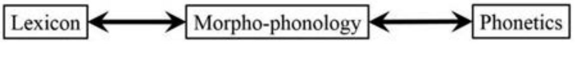


Figure 4.1 Modular grammatical architecture

4.2.2 Stratal OT: an overview

As noted above, Stratal OT is a constraint-based model of phonology that draws upon research carried out in the LPM programme (Kiparsky 1982; Mohanan 1986). Unlike rule-based theories, OT is characterised by direct input→output computation of phonological structures. By way of a recap from Iosad (this volume), Figure 4.2 below illustrates the operation of the computational mechanism assumed in classic OT (adapted from McCarthy 2002: section 1.1.4). Under this model, input structures are evaluated for phonological well-formedness (i.e. optimality) in parallel, by a single pass through GEN and EVAL (see Krämer, this volume: section 2.1). Thus, all properties of output candidates are assessed simultaneously by a single, language-specific ranking of the constraints in CON.

In contrast to the classic model, Stratal OT synthesises optimality-theoretic evaluation with the phonological cycle inherited from LPM. As illustrated in Figure 4.3, the product of this synthesis is a constraint-based model of phonology in which candidate evaluation applies *recursively*, to increasingly large morpho-syntactic units. In the Stratal OT model proposed by Bermúdez-Otero (2007, 2011), the phonological component of the grammar comprises three derivational levels (or *strata*), namely the stem level, the word level and the phrase level (SL, WL, PL, respectively). Each level contains a stratum-specific OT grammar

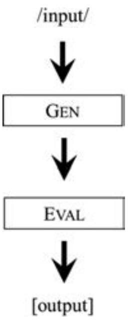


Figure 4.2 Computational procedure in classic OT

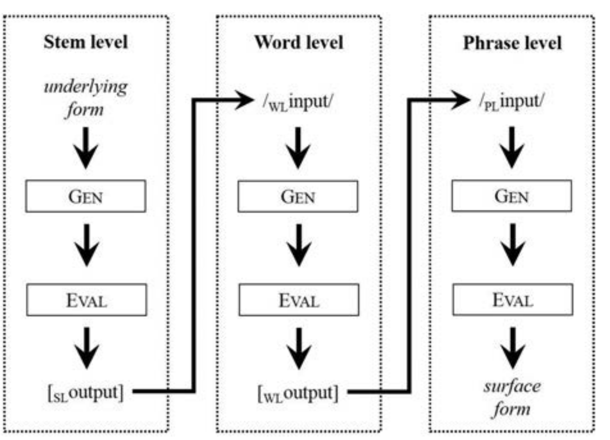


Figure 4.3 Computational procedure in Stratal OT

(i.e. a stratum-specific raking of the constraints in CON): the task of these level-specific grammars is to select an optimal output form from the series of candidates generated from the phonological input structures that are visible to the grammar at each level.

Morpho-phonological processes operate on three morphological constituents in Stratal OT, namely root, stem and word. As discussed in greater depth in Bermúdez-Otero (this volume), roots are uninflectable base units which do not define their own cyclic domain. Stems, by contrast, are free to undergo inflection and may be targets for phonological processes at the stem level. Words are fully inflected units and trigger phonological operations in the second, word-level stratum. Finally, the phrase-level phonology corresponds to the post-lexical stratum of LPM: categorical phrase-level processes therefore apply in maximal morpho-syntactic domains.

For example, consider the Standard Italian word *libro* 'book' which has the morphological structure $[\![_{\text{Word}}[\![_{\text{Stem}}\text{libɾ}_{\surd}\text{-o}_{\text{Th}}]\!]]\!]$. The bare root /libɾ/ cannot trigger any phonological processes; a morphological operation that supplies a class-marking suffix (here the theme vowel /-o/) is necessary to convert /libɾ/ into a stem (i.e. /libɾo/) which the stem-level phonology can act upon. Submitting this structure to the phonology subjects it, firstly, to any phonological processes that are triggered by evaluation of possible output forms against a ranking of markedness and faithfulness constraints at the stem level. Note that unlike LPM, Stratal OT does not impose the strict cyclicity condition (Bermúdez-Otero & McMahon 2006; Collie 2007: section 1.4): all phonological material that is visible to the grammar in a derivational stratum is a potential target for phonological processes that apply within that stratum.

The optimal output form generated by the stem-level grammar is then taken as the input to the next derivational level. The word-level phonology is responsible for generating the faithful mapping of /libɾo/→[libɾo]. Likewise, the constraint ranking at the phrase level ensures that the optimal output of the word-level phonology maps faithfully to [libɾo] in any phrasal contexts that it may occur in (e.g. /libɾo#peɾikoloso/→[libɾopeɾikolozo] 'dangerous book').[2]

In cases of affixation, the derivation is more complex. (1) below illustrates the derivation of *libretto* 'small book, (operatic) libretto' which is a diminutive form of *libro*. Here, we assume that the diminutive suffix /-etː-o/ attaches to the stem [libɾo] at the word level. Hence, the morphological structure of *libretto* is $[\![_{\text{Word}}[\![_{\text{Stem}}\text{libɾ-o}]\!]\text{-etː-o}]\!]$.

(1)

a. SL input: /libɾ-o/		*HIATUS	MAX-V_2	MAX
i.	[libɾo] ☜			
ii.	[libɾ]			*!
iii.	[liro]			*!

b. WL input: /libɾo/		*HIATUS	MAX-V_2	MAX
i.	[libɾo] ☜			
ii.	[libɾ]			*!
iii.	[liro]			*!

c. WL input: /libɾo-etː-o/		*HIATUS	MAX-V_2	MAX
i.	[libɾoetːo]	*!		
ii.	[libɾetːo] ☜			*
iii.	[libɾotːo]		*!	*

In the citation form [libɾo], segmental deletion is militated against both at the stem level (1-a) and at the word level (1-b) by the general MAXIMALITY constraint. This constraint demands that all segments present in the input must also be present in the output. By contrast, the stem-final [o] that occurs in [libɾo] is a target for deletion in [libɾetːo]. As the tableaux in (1) show, the phonology first evaluates the innermost bracket in ⟦$_{\text{Word}}$⟦$_{\text{Stem}}$libɾo⟧-etː-o⟧ against the stem-level constraint ranking: the faithful candidate [$_{\text{SL}}$libɾo] is selected as the winner. The second cycle then evaluates all phonological material contained within the outer bracket. In other words, the input to the word-level cycle is the output of the stem cycle plus the suffix /-etː-o/. A fully faithful mapping of /libɾo-etː-o/ to *[libɾoeto] incurs a violation of the superordinate *HIATUS constraint, which militates against [VV] sequences in the output. The remaining constraints in the ranking must therefore decide which of the offending vowels will be deleted. The higher ranking of a constraint enforcing the preservation of /V_2/ in a /V_1V_2/ string (MAX-V_2) relative to general MAX causes form (1-c-ii) with deletion of word-medial morpheme-final /o/ to be selected as the winner.[3]

4.2.3 Phrasal phenomena in Stratal OT

In addition to phonological processes that are driven by morphological operations like diminution, Stratal OT is particularly well suited to accounting for phonological processes that are triggered by the concatenation of words into a phrase. For example, (2) below gives the derivation of *libro pericoloso* for Tuscan Italian.[4] As mentioned, GT is a variable lenition phenomenon that causes intervocalic stops to spirantise, either word internally or across word boundaries. It also applies to post-vocalic stops that are immediately followed by glide or a liquid, hence /libɾo/→[liβɾo] (Sorianello 2002: 29; Marotta 2008: 242–243). Existing descriptions of GT state, however, that spirantisation does not occur in sonorant+stop sequences, hence /aɾt-e/→[aɾte], *[aɾθe] 'art'.

(2)

Domain structure:	⟦$_{\text{Phr}}$⟦$_{\text{Wrd}}$⟦$_{\text{Stm}}$libɾ$_{\surd}$-o$_{\text{Th}}$⟧⟧	⟦$_{\text{Wrd}}$⟦$_{\text{Stm}}$⟦$_{\text{Stm}}$peɾikol$_{\surd}$-o$_{\text{Th}}$⟧-os-o⟧⟧⟧
Stem level:	/libɾ-o/ ↓ [libɾo]	/peɾikol-o-os-o/ ↓ [peɾikolozo]
Word level:	/libɾo/ ↓ [libɾo]	/peɾikolozo/ ↓ [peɾikolozo]
Phrase level:	/libɾo # peɾikolozo/ ↓ [liβɾoɸeɾixolozo]	

As discussed in Iosad's chapter (this volume), categorical spirantisation can be modelled straightforwardly in OT as an interaction of a positional markedness constraint penalising intervocalic stops in the output with a faithfulness constraint demanding preservation of input manner-of-articulation features on the surface.[5] Candidate (b) in (4) below wins because it incurs no violations of the top-ranked markedness constraint. The cost of maximally respecting the demands of this constraint is violation of the lower ranked faithfulness constraint.

(3)

a. *[VC$_{[-\text{cont}]}${G$_0$,L$_0$}V]: assign one violation mark for every intervocalic stop, intervocalic stop+glide sequence or intervocalic stop+liquid sequence in the output.

b. Ident-MoA: assign one violation mark for every unfaithful mapping of an input manner-of-articulation feature in the output.

(4)

Input: ⟦$_{\text{Phr}}$ libroperikolozo⟧	*[VC$_{[-\text{cont}]}${G$_0$,L$_0$}V]	Ident-MoA
a. [$_{\text{PL}}$libroperikolozo]	*!**	
b. [$_{\text{PL}}$liβroɸerixolozo] ☜		***

Whilst this sketch-analysis captures the essence of the Tuscan pattern, it is incomplete in a number of ways. Firstly, GT varies across dialects. As discussed in greater detail in section 4.3 below, eastern varieties of Tuscan Italian typically spirantise to a greater extent that western varieties, and areal variability may be also compounded by style-dependent variability. Secondly, whilst there is evidence that GT can be a categorical phenomenon, it may also vary gradiently in theoretically interesting ways. Indeed, the picture that emerges from consideration of currently available data is that GT simultaneously has both phonological and phonetic characteristics.

4.3 A closer look at *Gorgia Toscana*

As already noted, GT may apply both word-medially and across word boundaries. An underlying stop is realised as a continuant either when it occurs in intervocalic position, or when it is preceded by a vowel and followed by a sonorant+vowel sequence (i.e. /V__V/ or /V__C$_{[+\text{son}]}$V/ environments).[6] Crucially, GT is not sensitive to morphological structure: in this sense, it is a purely phonotactic phenomenon.

Yet beyond these general statements, it is also clear that GT is by no means a homogeneous phenomenon. Existing research has confirmed that use of GT is very much dependent upon dialectal factors. For example, Hall (1949) discusses geographical variation in patterns of /p, t, k/-spirantisation recorded in linguistic survey material.[7] Hall notes that spirantisation of /k/ is observed, to some extent, across almost all of Tuscany. Spirantised variants of /t/, by contrast, appear to be quite a bit more common in Eastern Tuscany than in the West. Furthermore, /p/-spirantisation is almost exclusively confined to Eastern Tuscany with only scant attestation in Western Tuscan.

From a historical phonological perspective, these findings are particularly interesting. Synchronic dialectal variation of the sort uncovered by Hall is precisely what we should expect to observe under a situation of *diachronic rule generalisation* (Vennemann 1978: 260–261; Bermúdez-Otero 2013: section 3.1; Ramsammy 2015: section 4). In this scenario, the historical innovation that gave rise to GT is frication of intervocalic /k/. In the first instance, it is likely that this innovation took hold in the city of Florence and then spread outwards throughout the rest of Tuscany (see Marotta 2008: 240–241; Montemagni, Wieling, de Jonge & Nerbonne 2013: 157). In a later phase of innovation, intervocalic spirantisation then came to affect the coronal stops as well as the velars; and later still, spirantisation eventually affected the labial stops too. This development is illustrated in Figure 4.4 below.[8]

At timepoint 1 in Figure 4.4, the spirantisation rule targets only intervocalic dorsal stops. By timepoint 2, however, the spirantisation rule has become less restrictive and thus targets both intervocalic coronal stops as well as intervocalic dorsal stops. At timepoint 3, the rule has become even less restrictive and targets all intervocalic stops, irrespective of place of articulation. Crucially, this trajectory of change mirrors the dialect geographical patterns that Hall describes. In the most conservative western dialects, only the oldest pattern of /k/ spirantisation occurs robustly: it seems that the younger pattern of /t/-spirantisation had not yet had time to work its way fully westwards by the time the *AIS* data were collected (i.e. in the 1930s and 40s). Thus, there were presumably speakers in the west of Tuscany at this time who regularly spirantised intervocalic /k/, but who did not necessarily use /t/-spirantisation with the same degree of regularity. Likewise, the youngest version of GT in which dorsal, coronal and labial stops are targets for spirantisation remains geographically restricted to the linguistically innovative eastern regions of Tuscany. This means that a high proportion of speakers in the more conservative west did not spirantise intervocalic /p/ at the time the *AIS* data were recorded.[9] Nevertheless, it is also worth remembering that outward spread of GT had not reached completion at this point: for example, Gianelli & Savoia (1978: 24) have noted accelerated spread of GT into more peripheral areas since collection of the *AIS* data.

Understanding the variability of spirantisation processes that make up GT as the result of diachronic rule generalisation allows us to account for the fact that not all Tuscan speakers spirantise stops contrasting in place-of-articulation features with the same regularity. However, the trajectory of change outlined in Figure 4.4 provides no insights regarding a second important fact about GT, namely that the output of spirantisation also varies. In fact, the *AIS* data very clearly show that spirantisation can have different outcomes (Hall 1949: 65). Recorded realisations of /t/, for example, span a continuum from fully occluded [t] to pre-aspirated [^{h}t], post-aspirated [t^{h}] and fully fricated [θ]. Likewise, /p/ may be realised as a plain stop, an aspirated stop or as a full fricative;[10] and /k/ occurs as [x] or [h], and may also be fully elided.

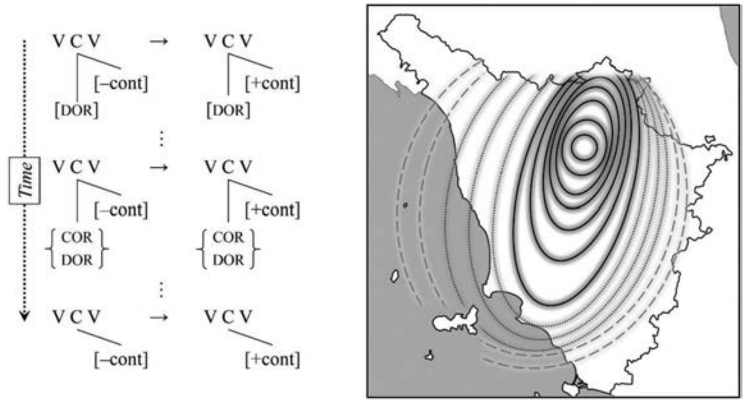
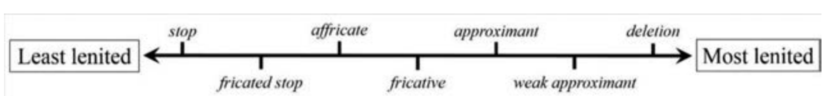


Figure 4.4 Diachronic generalisation (left) and geographic spread (right) of spirantisation. Map: solid lines = spirantisation of all stops, dotted lines = spirantisation of {COR, DOR} stops, dashed lines = spirantisation of [DOR] stops only

Later studies have also highlighted the gradient nature of GT. In particular, Gianelli & Savoia (1978) stress the importance of considering multiple interacting factors when attempting to account for the variability of GT. In addition to socio-economic status, educational background and age, speaking style is shown to play an important role (see also Kirchner 2004). For example, fricated realisations (i.e. [ɸ, θ, x]) are typical of slower, more accurate natural speech, and partially voiced realisations (i.e. [ɸ̬, θ̬, x̬]) are also observed at faster speaking rates. The most lenited realisations of /p, t, k/ (including deletions) occur in casual, emotional speech.

In addition, Villafaña-Dalcher (2008) presents an analysis of GT based on experimental data from six Florentine speakers. According to acoustic properties of realisations of /p, b, t, d, k, g/ in the relevant weakening contexts, Villafaña-Dalcher observes that the output of GT varies along the following lenition scale (cf. Honeybone 2012: section 3.3 on stop lenition in Liverpool English; see Figure 4.5 below).[11]

Although spirantisation is clearly variable, one interpretation of the data given in Figure 4.5 is that GT has both categorical and gradient characteristics. Firstly, a clear bimodality in the data can be observed in the barchart shown at the bottom of Figure 4.5. One peak represents a cluster of realisations towards the left of the x-axis (i.e. stops and fricated stops), and the second peak is formed by a cluster of fricative realisations. Secondly, note that this pattern holds across the three places of articulation. The fact that 85% of /k/-tokens are realised as fricatives or approximants suggests the operation of a discrete phonological process that lenites intervocalic stops. This process has not applied in the 18% of /k/-tokens



	N	stop (%)	fricated stop (%)	affricate (%)	fricative (%)	approx. (%)
/k/	231	4	14	0	62	23
/t/	232	27	22	4	47	0
/p/	197	25	7	7	60	1

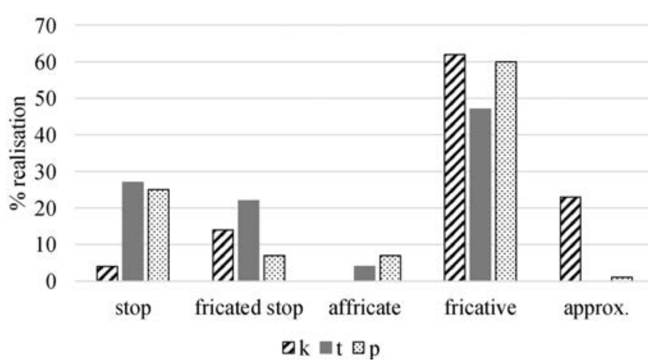


Figure 4.5 Lenition scale for GT (top) and lenition statistics for voiceless stops (bottom) (adapted from Villafaña-Dalcher 2008)

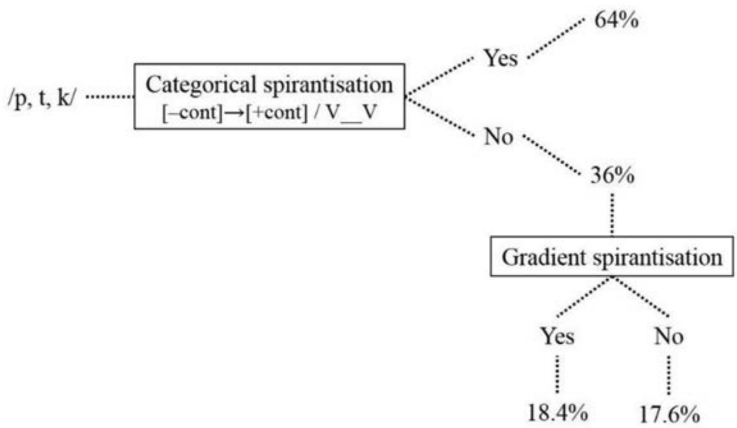


Figure 4.6 Categorical vs gradient spirantisation of intervocalic voiceless stops based on data from Villafaña-Dalcher (2008)

that are realised as stops; yet 14% of those tokens also show evidence of marginal levels of frication which in turn suggests the operation of a more gradient version of spirantisation. With regard to the /p/-realisations, 61% of tokens display the effects of what could arguably be a discrete spirantisation process, whereas the remaining 39% of tokens retain some degree of stop-like labial occlusion. 25% of these realisations are fully occluded plosives, whereas 14% display evidence of gradient spirantisation. Similarly for /t/, 47% of tokens are categorically spirantised whereas some degree of stop-like occlusion is observed in 53% of tokens. Of these 53%, 26% display some gradient lenition and 27% are fully occluded stops.[12] As illustrated by Figure 4.6 below, this interpretation of the facts means that discrete spirantisation – i.e. a categorical phonological process that converts underlying stops into continuants – applies in 64% of cases. By contrast, 18.4% of voiceless stops display traces of gradient spirantisation and 17.6% show no evidence of spirantisation whatsoever.

The general picture that emerges from these studies, therefore, is that GT is a variable phenomenon. However, the fact that style affects the degree of lenition at the intraspeaker level implies that the outcome of spirantisation is under the cognitive control of individual speakers. In this sense, spirantisation is not an automatic phenomenon (although automatic effects of gestural and aerodynamic interpolation may also play a role: see sections 4.4–4.6 below). In proposing any analysis of GT, we must therefore be mindful that it bears the hallmarks of an interface phenomenon. This is to say that the variable output of spirantisation may depend on phonological operations, phonetic implementational mechanisms and the idiosyncratic speech habits of individual speakers. Accordingly, GT provides excellent case-study material for examining how synchronic phonology–phonetics interactions can be modelled and accounted for theoretically, and for illustrating the role of phonology–phonetics interactions in processes of phonological change.

4.4 *Gorgia Toscana*: the view from the life cycle

Assuming the same grammatical architecture employed in Stratal OT for analysing synchronic morpho-phonological alternations, the *life cycle of phonological processes* (Bermúdez-Otero 2007, 2012; Ramsammy 2015) is a theory of phonological change capable of modelling the development of phonological phenomena over time. The life cycle is based on two core claims: firstly, that the locus of phonological innovation is the phonology–phonetics interface, and secondly, that categorical phonological processes become increasingly integrated with morpho-syntax as they age. As we shall see, the life cycle affords us important insights in attempting to reconstruct the pathways of change that have led to the synchronic patterns of spirantisation in Tuscan Italian.

It has long been understood that epiphenomenal phonetic events that occur in speech because of physiological or perceptual limitations of speakers and listeners can, over time, develop into fully fledged phonological processes. For example, there is tendency for f_0 to drop at the point of transition between a voiced stop and a following vowel and for f_0 to rise at the Onset of a vowel following release of a voiceless stop. This occurs because of aerodynamic repercussions of coordinating lingual gestures with changes in laryngeal articulatory configurations. Such marginal fluctuations in f_0 in stop+vowel sequences are, to some extent, automatic, and may escape perceptual identification. However, in a number of languages, the phonetic tendency for f_0 to rise or fall depending upon the voicing of a preceding consonant has resulted in tonogenesis. This is to say that epiphenomenal f_0 rises and falls in CV transitions have phonologised diachronically into a contrast between high and low tone. In some languages (e.g. certain dialects of Kammu), tonogenesis has also caused loss of the voicing contrast on stops that originally conditioned the gradient tonal contours (Ohala 1993: 239–240; Kingston 2011: section 2.3). In other words, phonologisation of tonal epiphenomena in Kammu varieties has led to the replacement of an original laryngeal contrast (voiced vs voiceless stops) with a tonal one (high vs low tone).

Under the life cycle, the emergence of segmental innovations works in the same way as tonogenesis. Understanding what we might analogously dub ‘spirantogenesis’ for Tuscan Italian therefore depends on identifying an automatic phonetic effect that, over time, could have developed into a phonological rule of stop spirantisation. We cannot say with certainty what originally conditioned the emergence of GT: recall that the spirantisation of /k/ is known to have occurred in Tuscan speech since the 16th century. Nevertheless, in attempting to reconstruct a trajectory of change that plausibly could lead to the establishment of a new phonological rule of stop spirantisation, a reasonable working hypothesis is that gestural undershoot may have played a role. As shown in the articulatory window diagram (Keating 1990) in Figure 4.7 below, production of a stop like [k] requires sealing off the oral tract by

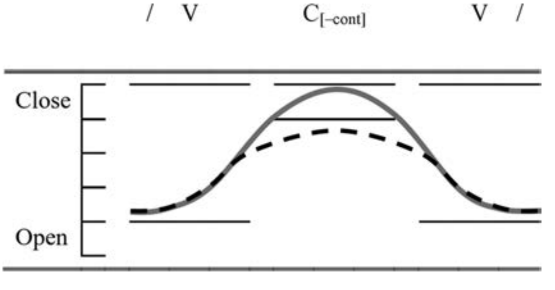


Figure 4.7 Gestural implementation of a /$VC_{[-cont]}V$/ sequence. The solid contour schematises accurate achievement of all articulatory targets including the formation of a complete oral seal for the realisation of the intervocalic stop. The dashed contour represents a realisation of the same /$VC_{[-cont]}V$/ sequence in which the target for the intervocalic stop is undershot: this results in an incomplete articulatory seal and the generation of aerodynamic turbulence across the partial closure

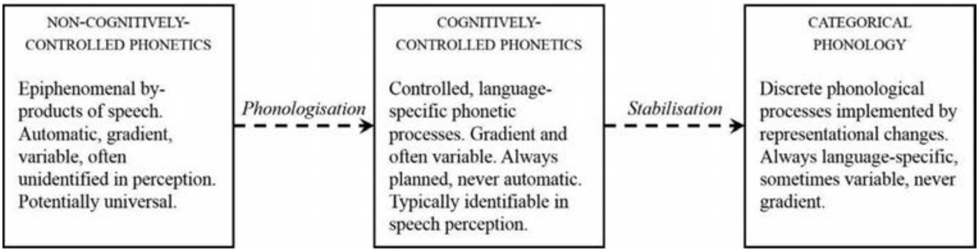


Figure 4.8 Phonologisation and stabilisation in the life cycle of phonological processes

the formation of a close occlusion between the tongue dorsum and the velum. However, if the target for full oral closure is undershot,[13] then air particles can travel through the oral cavity during the closure phase of the stop. Crucially, this type of undershoot has acoustic as well as aerodynamic consequences. Rather than the production of a plosive with crisp closure transitions and release burst, incomplete occlusion causes partial frication of the stop as air particles pass through the incomplete seal under pressure.

Cases of lenition resulting from undershoot of stop-closure targets have been well documented (at least in the early stages of sound change; see Simpson 2001: section 3). Thus, it is plausible that gestural undershoot of the type illustrated in Figure 4.7 above may have led to /k/ being realised with some degree of frication in the historical varieties of Tuscan in which GT subsequently developed (cf. Marotta 2008: 250). It is also likely that these sorts of undershoot phenomena should occur more frequently in more casual registers when speakers are paying less attention to careful articulation of speech sounds. This is important since it is in the most relaxed speech styles that the most extreme synchronic lenitions of intervocalic stops in Tuscan Italian are observed (Gianelli & Savoia 1978, 1979–80). Yet if we are correct in our hypothesis that gestural undershoot is the automatic phonetic phenomenon that eventually gave rise to GT, then it follows that partial stop frication underwent phonologisation in Tuscan Italian in the same way that gradient tonal contours have phonologised in languages like Kammu. Phonologisation has a specific definition in the life cycle framework: it occurs when an uncontrolled, automatic by-product of speech is reanalysed over time as a phonetic process whose occurrence is under the cognitive control of the speaker (see Ramsammy 2015: section 2).[14] Following phonologisation, the life cycle also predicts that gradient, but cognitively-controlled, phonetic phenomena can then develop into categorical phonological rules. This second phase of change is termed *stabilisation*.

Before addressing the question of how phonologisation and stabilisation can be modelled in an OT framework, we must first address a broader question: namely, how do phonological and phonetic structures interact in constraint-based models of grammar? Section 4.5 explores this question with a focus on a constraint-based framework that makes specific claims about phonology–phonetics interactions. Section 4.6 then considers how this model can be applied to the phonologisation and stabilisation processes schematised in Figure 4.8.

4.5 The phonology–phonetics interface in constraint-based grammar

Although many (if not most) analyses of phonological phenomena in OT assume that the product of phonological computation (i.e. the winning output candidate) is fed into a phonetic implementation module, the mechanisms responsible for implementing the transduction process are rarely given much attention. Similarly, whilst the roles of phonologisation and stabilisation in processes of phonological change have been emphasised in existing research within the life cycle paradigm (e.g. see Bermúdez-Otero 2007, 2012), there has been little detailed discussion of the processes involved in these developments. Nevertheless, constraint-based phonology provides a framework for modelling interactions between the phonological and phonetic modules of grammar. In *Bidirectional Phonology and Phonetics* (henceforth BiPhon: Boersma 2007, 2009; Boersma & Hamann 2008; Hamann 2010), constraint interaction is the mechanism behind all phonological computation, both in terms of discrete phonological operations and gradient phonetic implementational operations. Figure 4.9 below presents a schematisation of the BiPhon model (adapted from Boersma 2009: section 1, 2007: 2031–2032). As a tool for explaining synchronic phonological phenomena, the advantage of this model is its ability to model both production and perception. As shown, production happens by feed-forward computation. The first stage of this process is retrieval of lexically stored forms (here, <morphemes>). Lexical structures are fed into the phonological module which derives SRs from URs by the OT evaluation procedure. Optimal SRs – i.e. winning output forms – are then submitted to phonetic implementation. Here, an optimal Auditory Form is generated by evaluation of competing candidates against a ranking of *cue constraints*. Cue constraints enforce well-formed mappings of discrete phonological representations to language-specific auditory targets. The winning Auditory Form is then fed forward to the gestural planning module which is responsible for calculating an articulatory plan capable of generating speech that maximally reflects the auditory-acoustic properties of the auditory prototype.[15]

In perception, by contrast, the hearer does not have access to the speaker's Articulatory Form: thus, the Auditory Form is the input to the speech-comprehension process. As shown, the hearer's task is to map the continuous acoustic information contained in the Auditory Form to a discrete phonological surface form. Knowledge of the relative ranking of structural and faithfulness constraints in the language guides the listener in the word recognition process which entails associating the SR with a UR and, in turn, with retrieval of the correct lexical forms.

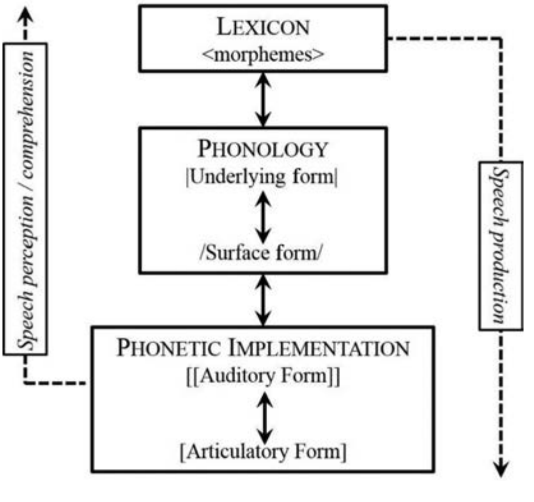


Figure 4.9 The BiPhon model

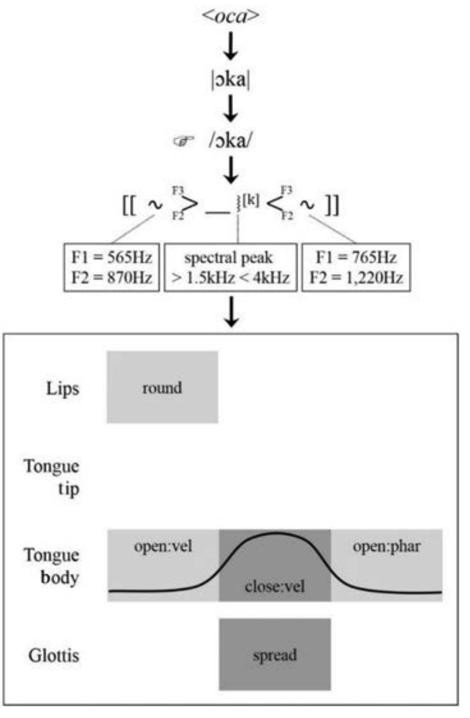
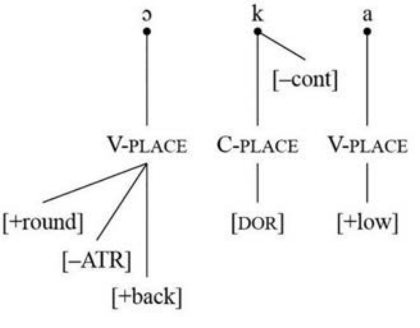


Figure 4.10 Production of *oca*

Returning to speech production, Figure 4.10 above illustrates the production process using the example of *oca* 'goose' as it would be pronounced in Standard Italian. Retrieval of the lexical item <*oca*> instantiates submission of the underlying form |ɔka| to the grammar. Assuming a stratified phonology (e.g. as shown in Figure 4.3), /ɔka/ is selected as the optimal SR by the cumulative evaluation of three iterations of GEN and EVAL at stem, word and phrase levels, respectively. The winning surface form, /ɔka/, is then fed forward to phonetic implementation. Note here that intervocalic stop spirantisation is not part of the phonology of Standard Italian: thus, the dialect-specific ranking of faithfulness constraints ensures that *[ɔxa] is never deemed more harmonic than [ɔka] at any stage in the derivation.

Let us assume that the phonological output /ɔka/ has the following representation:

(5)



At this point, the phonological surface structure undergoes transduction into an auditory object. In BiPhon, as noted, the mechanism responsible for this conversion is a system of ranked cue constraints. In (6), possible Auditory Forms generated from the surface form shown in (5) are submitted for evaluation. As with structural and faithfulness constraints in the categorical phonological module, the ranking of cue constraints is determined in acquisition on a language-specific basis (see Boersma & Hamann 2008).[16] Here, cue constraint (6i) requires that segments bearing a [−cont] feature value on the surface should be realised with a period of silence (corresponding to the closure phase in stop articulations). This constraint is decisive in eliminating candidate (6b) from the running (silence is represented as [[__]] in (6) and Figure 4.10). Similarly, candidates (6a–d) and (6f) satisfy a second constraint, (6ii), demanding the mapping of the surface structure /k/ (i.e. [−cont, DOR]) onto an Auditory Form containing a stop burst (represented here by [[⦚]]) with the specific acoustic properties of dorso-velar release, [[⦚[k]]] (i.e. a spectral prominence between 1.5 and 4kHz).[17] The cue constraint for velar burst is decisive in eliminating candidate (6e) in which the stop burst has the expected acoustic characteristics of labial release (e.g. a spectral peak below 1kHz). Candidate (6f) is eliminated due to its violation of cue constraints (6v–vi) which militate against auditory objects lacking convergences of F2 and F3 (i.e. a 'velar pinch') during the approach transition into, and the release transition from, the intervocalic stop.[18]

(6)

SR: /ɔka/	i. [−cont]↔[[__]]	ii. /k/-RELEASE↔[[⦚[k]]]	iii. /ɔ/↔[[F1=565, F2=870]]	iv. /a/↔[[F1=765, F2=1,220]]	v. /k/-CLOSURE↔ [[>]]	vi. /k/-RELEASE↔ [[<]]	vii. /ɔ/↔[[F1=440, F2=770]]	viii. /a/↔[[F1=540, F2=1,870]]
a. [[ɔ >__⦚[k] < a]] ☜							*	
b. [[ɔ >⦚[k] < a]]	*!						*	*
c. [[o >__⦚[k] < a]]			*!				*	
d. [[ɔ >__⦚[k] < ɛ]]				*!				*
e. [[ɔ >__⦚[p] < a]]		*!					*	*
f. [[ɔ =__⦚[k] = a]]					*!	*!	*	*

The mapping of the vowels in /ɔka/ onto language-specific auditory targets is also determined by cue constraints. The high ranked constraint in (6iii) requires that surface /ɔ/ should be associated with a periodic soundwave (designated by [[∿]] in Figure 4.10) in which F1 occurs at 565Hz and F2 at 870Hz.[19] Likewise, (6iv) militates against any auditory mapping in which surface /a/ is not associated with a periodic soundwave with an F1 of 765Hz and an F2 of 1,220Hz. These constraints are responsible for the elimination of candidates (6c) and (6d), respectively. In (6c), surface /ɔ/ maps to [o] in the Auditory Form. This candidate satisfies the demands of a low ranked constraint (6vii) demanding association of /ɔ/ with a periodic wave in which F1 occurs at 440Hz and F2 at 770Hz (i.e. the expected values for realisation of close-mid /o/ in Standard Italian). However, it crucially violates the higher ranked cue constraint (6iii) demanding that /ɔ/ should be associated with an auditory object bearing the language-specific acoustic properties of open-mid /ɔ/. Similarly, candidate (6d) is deemed suboptimal given than surface /a/ maps to an auditory structure in which F1 occurs at 540Hz and F2 at 1,870Hz. These values are acceptable for realisation of the vowel /ɛ/ in Standard Italian, but not for /a/. Accordingly, (6d) loses out to (6a) because of its fatal violation of (6iv). This occurs despite the fact that (6a) violates a low ranked cue constraint, (6viii), which requires mapping of /a/ to vocalic structure with formant values that would be typical for a realisation of the vowel [ɛ] in Standard Italian.[20]

The next stage in the production process is the mapping of the winning Auditory Form onto an Articulatory Form. Recall from Figure 4.9 that the Articulatory Form is an articulatory plan whose implementation should facilitate maximal achievement of the auditory targets specified in the Auditory Form. The *gestural score* (see Browman & Goldstein 1986, 1990, 1992) shown in the lower panel of Figure 4.10 is the outcome of mapping [[ɔ >__⌇[k] < a]] onto gesture-based articulatory plan. Observe here that [ɔ] involves the coordination of two articulatory movements, namely lip rounding and the formation of an open constriction between the tongue dorsum and the velum. By contrast, there is no lip-rounding target for production of post-consonantal /a/: articulation of this vowel instead requires the realisation of a target for an open linguo-pharyngeal constriction. Whilst vowel articulations involve relatively wide oral constrictions, note that articulation of /k/ requires a close dorso-velar constriction. This gesture must also be coordinated with a target for glottal opening in order to prevent the occurrence of vocal fold vibration and, in turn, the generation of perceptible voicing.

Other things being equal, realisation of the tongue body contour represented by the solid black curve in Figure 4.10 in coordination with the non-lingual gestures shown is expected to result in generation of speech in which the crucial acoustic information specified in the Auditory Form [[ɔ >__⌇[k] < a]] is faithfully represented. Following formation of articulatory gestures for [ɔ], the crucial transitional cues necessary to make [k] recoverable from the speech signal are produced as the tongue body interpolates between the vowel constriction target and the stop constriction target. A period of silence is manifested as cessation of modal voicing, facilitated by vocal fold abduction: this coincides with maximal displacement of the tongue dorsum and the formation of an articulatory seal against the velum. Release of this occlusion has various acoustic consequences. Firstly, a transient burst with a dorso-velar acoustic colouring is heard as the raised dorsum separates from the velum. Secondly, release of muscular tension responsible for stiffening of the vocal folds and glottal widening during the realisation of [k] results in the re-initiation of vocal fold vibration and, thus, periodic voicing. Thirdly, gradual displacement of the tongue dorsum towards the articulatory target for [a] causes a rise in the frequency of F3 and a depression of the frequency of F2: a 'release pinch' is therefore perceptible in the transitional periods between [k] and [a].

Whilst realisation of the gestural score shown in Figure 4.10 means that the auditory information specified in the Auditory Form selected by (6) should be present in the speech signal, we ought also to ask how the mapping between the Auditory Form and the Articulatory Form is managed. Can this process, too, be handled by constraint evaluation? In comparison to the large number of studies in which constraint-based models have been applied to account for categorical phonological phenomena in the languages of the world, potential applications of OT for dealing with gestural phonetic phenomena remains a relatively underexplored area of the theory. However, some have recognised this gap. For example, highlighting a need for any complete theory of phonology to include mechanisms for accounting for gestural phenomena,[21] Gafos (2002) argues that the temporal organisation of gestures can be modelled in a constraint-based framework by capitalising on and further developing constraints designed to enforce linear alignment of phonological structures in OT. Specifically, a set of *coordination constraints* are responsible for enforcing the alignment of *landmarks* defined upon contiguous gestures (e.g. centre of the gestural plateau, release offset etc.: see Gafos 2002: section 3 for full discussion; see also Bradley 2006 on vowel excrescence in Spanish stop+rhotic clusters).

Yet in addition to phenomena that originate in patterns of overlap between articulatory gestures, accounting for phenomena like GT in the BiPhon model requires a way of relating auditory targets to articulatory events that are not solely determined by the horizontal organisation of gestures. Figure 4.7 aims to capture the fact that GT is very much the sort of phenomenon that demands a theory of the mechanics of gestural planning in the vertical dimension (i.e. degree of occlusion).

(7) below presents a preliminary sketch of how mapping between Auditory Form and Articulatory Form might be implemented by constraint interaction. The input here is the Auditory Form selected as the winner by the cue constraints in (6). From this input, a number of possible gestural mappings are generated: note here that output candidates are listed in abbreviated form, where candidate (a), [ɔka], for example, is shorthand for the gestural score in Figure 4.10. The articulatory constraints here assign violations in the same manner as the cue constraints in (6): that is, violations are assigned to mappings between Auditory and Articulatory Forms that do not meet the specific demands of individual articulatory constraints.

(7)

AudForm: [[ɔ >__⸾[k] < a]]	i. [[ɔ]]→ Lips:Round	ii. [[ɔ]]→ TB:Vel	iii. [[a]]→ Lips: ∅	iv. [[a]]→ TB:Phar	v. [[> __⸾[k] <]]→ TB:Vel	vi. [[__]]→ Close Occl	vii. [[⸾]]→Crisp Release	viii. [[ɔ]] → Lips: ∅	ix. [[a]]→ Lips:Round	x. [[__]]→ Near Close Occl	xi. [[__]]→ Open Occl	xii. [[⸾]]→Slow Release
a. [ɔka] ☜								*	*	*	*	*
b. [ʌkœ]	*!		*!							*	*	*
c. [œkæ]		*!		*!				*	*	*	*	*
d. [ɔta]					*!			*	*	*	*	*
e. [ɔk̚‿ʔa]							*!	*	*	*	*	*
f. [ɔkˣa]							*!	*	*	*	*	
g. [ɔxa]						*!	*!	*	*		*	
h. [ɔɰ̊ a]						*!	*!	*	*	*		

For example, constraint (7i) penalises gestural mappings of the auditory object [[ɔ]] which do not include a target for lip rounding. Likewise, (7iii) forbids mappings of [[a]] in which a target for lip rounding is present in the gestural plan. These constraints therefore cause the elimination of candidate (7b) in which [ʌ] (< [[ɔ]]) lacks a lip-rounding target and [œ] (< [[a]]) has one. Constraints (7ii) and (7iv) work similarly: (7ii) requires that the articulatory plan for [[ɔ]] should include a target for lingo-velar constriction, and (7iv) requires that [[a]] should be realised with linguo-pharyngeal constriction. Violation of these constraints eliminates candidate (7c) due to the fact that [œ] is not articulated with a dorso-velar gesture and the primary constriction target for [æ] is not pharyngeal. Despite being eliminated by top-ranked constraints, observe that candidate (7b) respects the demands of lower ranked constraints (7viii) and (7ix) which require different mappings between auditory and articulatory objects from (7i) and (7iii).[22]

In addition, (7v, 7vi, 7vii) constrain the gestural mapping of the auditory cues for /k/. Constraint (7v) requires that the auditory targets in [[>__⦚[k] <]] must be achieved through the realisation of a dorso-velar constriction: observe that this constraint is decisive in eliminating candidate (7d) in which [t] represents linguo-alveolar constriction. Constraint (7vi) militates against gestural mappings of auditory silence that do not include a target for full occlusion between articulators. Candidates (7g) and (7h) lose to (7a) given that production of the continuants, [x] and [ŭ̊ɻ], involves the realisation of a near-close constriction and an open constriction, respectively. Despite their violation of the superordinate (7vi), candidates (7g) and (7h) nevertheless incur no violation of the lower ranked constraints in (7x–xi), both of which permit mappings of acoustic silence to articulatory targets for incomplete oral constriction. Finally, constraint (7vii) requires that the burst cues in the Auditory Form should map to an articulatory target for crisp release of articulators. In other words, this constraint demands that stop-like constrictions must be released with a sufficiently rapid and large displacement of active articulations in order to facilitate plosion.[23] Constraint (7vii) therefore eliminates candidate (7e) in which [k] is unreleased and glottalised, candidate (7f) in which [k] has a fricated release and candidates (7g) and (7h) which display intervocalic fricatives the release of which does not cause burst-like transience. Note that candidates (7f–g) exhibiting fricated releases incur no violations of the low ranked constraint in (7xii) which permits mapping of the auditory burst to a gestural plan that incorporates a slower, smaller displacement of active articulators at the point of release than constraint (7vii).

Thus, maximal satisfaction of the articulatory constraints causes candidate (7a) to be selected as the winner. Although the evaluation mechanisms suggested in (6) and (7) are intended only as a preliminary sketch of how mapping between discrete phonological outputs and auditory and articulatory objects might obtain in BiPhon, they illustrate that there is potential for processes included in the production strategy shown in Figure 4.10 to be handled by optimality-theoretic constraint evaluation. Further research will be necessary to work out the finer details of how both intramodular and intermodular translations between different phonological and phonetic structures can best be modelled through the use of violable constraints (see Cavirani 2015: sections 6–7 for relevant discussion). Nevertheless, the main advantage of assuming an analysis of the sort sketched out above is that it puts us in a position to return to an outstanding question: how can the theoretical claims of the BiPhon and life cycle models be applied to help us develop an account of GT? This is the question to which we now turn in sections 4.6.

4.6 Modelling phonologisation and stabilisation

4.6.1 Gorgia Toscana *as a gradient process*

Returning to the life cycle and the question of the diachronic emergence of GT, consider Figure 4.11 below. The bold line represents the tongue body contour in the realisation of [ɔka] (i.e. phonetic production of <*oca*> in which the intervocalic stop is fully occluded, as in Figure 4.10). However, observe that if the displacement of the tongue dorsum away from the velum in the release of the close dorso-velar occlusion for [k] involves a slower, longer movement (indicated in Figure 4.11 by the dashed line), a short period of turbulent frication occurs in the transition between [k] and the following vowel. The result is a token of <*oca*> with a fricated release, i.e. [ɔk^{x}a]. Similarly, if the articulatory approach to the occlusion target for [k] is also relaxed in the same manner (indicated by the dotted line), the result is a stop realisation characterised acoustically by noisier, more aerodynamically turbulent closure transitions (i.e. [ɔ^{x}ka]).

To reiterate a point already made, it is important to bear in mind that we cannot be sure what factors originally conditioned the emergence of GT in Tuscan Italian. However, in line with the theoretical predictions of the life cycle, it is reasonable to hypothesise that there may have been a tendency for dorso-velar occlusion targets to be undershot with some degree of regularity in intervocalic contexts, perhaps most frequently in rapid, casual, unmonitored speech, in the diachronic stage preceding the emergence of GT. Such an articulatory tendency has an acoustic consequence: that is, there is a tendency for intervocalic stops to sound partially fricated in casual speech.

A crucial factor in the life cycle of spirantisation is how a tendency for variable, partial frication of intervocalic [k] may have been interpreted by children acquiring the phonology of Tuscan Italian in this time period. If the adult speech model that children are exposed to throughout phonological acquisition includes a sufficiently high number of tokens of fricated intervocalic /k/, then the gradient frication pattern that is automatic and perhaps even subconscious for adult speakers of Tuscan Italian may undergo a *change in status* (in the sense of Anderson 1981). A later generation of speakers may internalise the

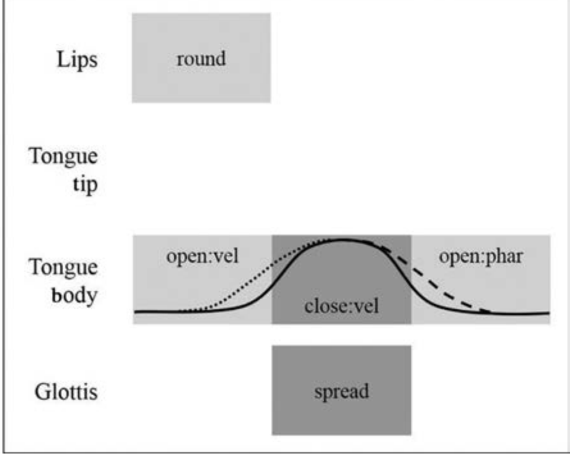


Figure 4.11 Gestural scores for [ɔka] (bold line), [ɔ^{x}ka] (dotted line) and [ɔk^{x}a] (dashed line)

non-cognitively-controlled stop frication pattern as a planned, stylistically dependent phonetic feature of Tuscan speech. In other words, a pattern of variation that is not under cognitive control in the speech of generation *n* may phonologise into a cognitively-controlled pattern in the speech of generation *n*+*x* by the pathway of change illustrated in Figure 4.8 above.

The question then arises of how a child of the younger generation is able to build a gradient phonetic process for articulatory relaxation of intervocalic stops into its grammar. Recall that, under the BiPhon model, hearers do not have access to the articulatory grammar of speakers: it is for this reason that the speech-comprehension process illustrated in Figure 4.9 begins with an auditory input, not an articulatory one. This same point is true of acquisition: children do not have direct access to the articulatory grammar of adult speakers; on the contrary, it is the speech signal (not articulatory plans) that forms the basis of phonological acquisition. It follows, therefore, that a child acquiring a grammar in which variable stop frication is a language-specific phonetic *innovation* arrives at a ranking of cue constraints that is different from adult speakers for whom stop frication is an automatic, epiphenomenal by-product of rapid speech. Thus, the adult grammar is one in which phonological surface structures such as that given in (5) map to Auditory Forms and Articulatory Forms containing targets for full stop occlusion in accordance with the demands of the constraint rankings shown in (6) and (7). The innovative grammar is one in which structures like (5) undergo mapping to a different Auditory Form.

(8)

SR: /ɔka/	i. [−cont]↔[[__]]	ii. [−cont]↔[[x]]	iii. /k/-Release↔[[⌇[k]]]	iv. /k/-Release↔[[[x]]]	v. /k/-Closure↔ [[>]]	vi. /k/-Release↔ [[<]]
a. /ɔka/↔[[ɔ >__⌇[k] < a]]		*		*		
b. /ɔka/↔ [[ɔ > __[x] a]]		*	*			*
c. /ɔka/↔ [[o x a]]	*		*		*	*

(8) above shows violations assigned by cue constraints relevant for mapping /k/ in the surface form /ɔka/ onto three alternative auditory objects. The constraints (8i–vi) are not presented here in a ranked hierarchy: the purpose of this tableau is to illustrate potential patterns of violation assigned by constraints relevant for auditory mapping /k/.[24] As we have already seen in (6), candidate (8a) respects the demands of cue constraints (8i) and (8iii), which require that /k/ should map to an Auditory Form containing a period of silence and a velar release burst, respectively. This candidate also incurs no violations of constraints requiring velar pinch transitions to be present in the Auditory Form: i.e. (8v) and (8vi). However, observe that (8a) does violate other cue constraints. (8ii), for example, is a cue constraint that permits non-continuants in the phonological output to correspond to an auditory target for aperiodic frication (represented here by [[x]]). Likewise, (8iv) penalises

auditory candidates lacking a fricated release (represented by superscript [[^{x}]]). Note that the alternative auditory candidates in (8b) and (8c), both of which lack a stop burst, incur violations of release-cue constraint (8iii) but not (8iv). Candidates like (8c) with an intervocalic fricative also violate (8v) and (8vi) which require the presence of clearly perceptible velar pinches in the auditory output. I assume here that the dorsum-raising gesture in the realisation of [x] is insufficiently large to cause clear narrowing of the acoustic space between F2 and F3: thus, for the same reason, candidate (8b) exhibiting a fricated release also incurs a violation of (8vi).

As in any OT grammar, it is the relative ranking of the cue constraints in (8) that determines which of the candidates shown will emerge as the winner. However, GT is a variable process: this means that each of the auditory candidates shown in (8) is a potential winner. As highlighted by the existing studies on GT discussed in section 4.3 (and in Gianelli & Savoia's work in particular), different patterns of lenition are observed in different registers. Thus, phonologisation of epiphenomenal stop frication must also involve a mechanism for ensuring *stylistic control* over the use of gradient lenition. In this connection, consider Figure 4.12 below.

Here, the idea is that implementation of a single phonological surface structure can yield planned, style-dependent patterns of variation. If we assume that acquisition of Tuscan Italian requires constructing a grammar that is capable of applying spirantisation in a stylistically appropriate way, then one way that this might be achieved is by register-specific rankings of cue constraints. In Figure 4.12 above, the output candidate in (a) is selected as the winner. This mapping applies in a ranking in which constraints demanding a canonical stop-like realisation of /k/ outrank those admitting some level of frication: crucially, [−cont]↔ [[__]], /k/-RELEASE↔[[⦚$^{[k]}$]] ≫ [−cont]↔[[x]], /k/-RELEASE↔[[$^{[x]}$]]. However, re-ranking these dependencies generates alternative patterns. Demotion of the constraint enforcing crisp plosion, for example, can generate mappings in which fricated realisations are deemed more harmonic: e.g. [ɔk^{x}a] ≻ [ɔka] (pattern 8b) if /k/-RELEASE↔[[⦚$^{[k]}$]] is ranked lower than /k/-RELEASE↔[[$^{[x]}$]]; and [ɔxa] (pattern 8c) wins out under a ranking in which [−cont]↔ [[x]] and /k/-RELEASE↔[[$^{[x]}$]] dominate [−cont]↔ [[__]] and /k/-RELEASE↔[[⦚$^{[k]}$]].

If we therefore make the assumption that speakers of Tuscan Italian construct style-specific co-phonologies like this (cf. Kager 1999: 404–405),[25] then the model shown in Figure 4.12 relates the variability of spirantisation to *alternative transduction strategies* that are responsible

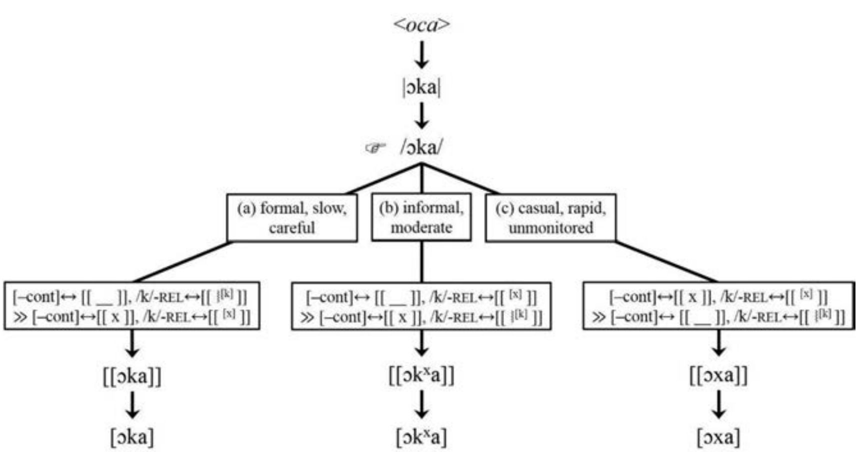


Figure 4.12 Style-dependent variation in cognitively-controlled gradient stop frication

for mapping winning phonological surface forms onto variable auditory representations. Moreover, in terms of dialectal patterns, this grammar potentially captures the spirantisation patterns used by the most conservative speakers in the most peripheral areas of Tuscany synchronically (i.e. with gradient spirantisation of /k/ but without any spirantisation of /t/ or /p/: cf. Figure 4.4). In diachronic terms, this grammar may also represent a historical stage of the dialects spoken in and around Florence in which spirantisation had not yet stabilised into a categorical process and in which gradient spirantisation was restricted to dorsal stops. If this is the case, then a final question to address is how spirantisation has expanded diachronically from a gradient phonetic process targeting intervocalic /k/ to one that, at least for some speakers, appears to involve categorical alternation between stops and fricatives at three places of articulation.

4.6.2 Gorgia Toscana *as a categorical process*

To recap from section 4.3, recall that GT bears all the hallmarks of a phonological process that has undergone diachronic rule generalisation in the sense of Vennemann (1978). Furthermore, the spirantisation patterns documented in the *AIS* show a very clear synchronic dialectal trend by which /k/-spirantisation occurs across Tuscany, whereas /t/-spirantisation is observed only in the central and eastern regions and /p/-spirantisation is restricted to eastern dialects. In this way, the synchronic dialect patterns mirror historical phonological innovation: the oldest pattern (i.e. /k/-spirantisation) has extended furthest geographically, whereas the youngest pattern (i.e. /p/-spirantisation) is synchronically confined to the area surrounding the urban centre of innovation in which GT first emerged (i.e. Florence).

In phonological terms, these developments present two questions. Firstly, how is rule generalisation – i.e. the pathway of change responsible for the diachronic expansion of Tuscan spirantisation – to be understood and modelled in a constraint-based framework? Secondly, how does the synchronic grammar of speakers of more conservative dialects differ from that of speakers of more advanced dialects?

To address the first question, we must consider how a variable, gradient process can be reinterpreted over time as a categorical phonological process. A possible route for this development is illustrated in Figure 4.13 below. Observe here that, at timepoint 1, the grammar

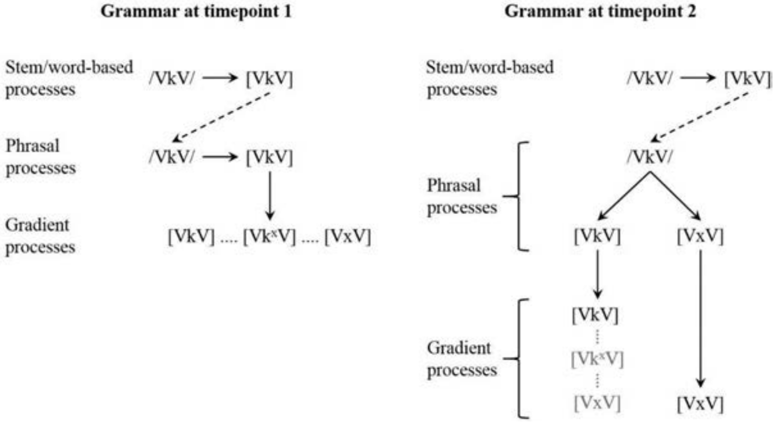


Figure 4.13 Stabilisation of /k/-spirantisation. Spirantisation is a cognitively-controlled gradient process in grammar 1 (left). In grammar 2 (right), gradient spirantisation has undergone stabilisation into a categorical phrasal process

of Tuscan Italian is identical to that shown in Figure 4.13. This means that the categorical phonology maps all instances of /k/ to [k]: gradient spirantisation may apply only in the phonetic implementational module. By timepoint 2, however, /k/-spirantisation has undergone stabilisation: at this point, its occurrence is under the control of the categorical phonology.

The key to this process is *restructuring* of the input to the phonetic module.[26] At timepoint 1, the input to the implementational process that generates gradient variable spirantisation is a single phonological surface structure, /k/. In the phase of change between timepoints 1 and 2, the pattern has been reanalysed: speakers/hearers reinterpret gradiently spirantised realisations of /k/ (represented in grammar 1 as the implementation scale, [k] . . . $[k^x]$ . . . [x]) as already being present in the output of the categorical phonology. This means that the phrasal grammar must be modified in order for it to be capable of generating surface forms displaying discrete but variable spirantisation. In other words, speakers acquiring Tuscan Italian at timepoint 2 construct a grammar in which the stratum-specific ranking of constraints at the phrase level is different from speakers who acquired the older grammar of Tuscan Italian at timepoint 1. The output of the categorical phonology – and, hence, the input to the phonetic implementational module – is therefore different for speakers who have acquired the innovative grammar from speakers who have the older grammar. But how does this change in the grammar take place? One way of thinking about the stabilisation process is to hypothesise that speakers acquiring grammar 2, confronted with the occurrence of spirantisation in the ambient language, construct new, language-specific markedness constraints that favour the use of spirantised realisations of /k/ in non-formal speaking contexts. Consider the following tableaux (cf. (2) in Iosad, this volume).

(9)

a. Phrasal grammar 1 (conservative)		
PL input: /ɔka/	IDENT-MoA	*$C_{[-\text{cont}]}$
i. [ɔka] ☜		*
ii. [ɔxa]	*!	

b. Phrasal grammar 2 (innovative, casual)			
PL input: /ɔka/	*$[VC_{[-\text{cont, DOR}]}\ \{G_0,L_0\}V]$	IDENT-MoA	*$C_{[-\text{cont}]}$
i. [ɔka]	*!		*
ii. [ɔxa] ☜		*	

c. Phrasal grammar 3 (innovative, formal)			
PL input: /ɔka/	IDENT-MoA	*$[VC_{[-\text{cont, DOR}]}\ \{G_0,L_0\}V]$	*$C_{[-\text{cont}]}$
i. [ɔka] ☜		*	*
ii. [ɔxa]	*!		

In grammar 1 in (9a), there is no categorical spirantisation: as illustrated in Figure 4.13, we assume that spirantisation can occur, but only as a gradient process. Accordingly, there is no constraint in the phrasal grammar that can generate continuants in the output from non-continuants in the input. In grammar 2, however, speakers have reinterpreted the gradient pattern as a categorical one: discrete spirantisation – i.e. changing an input [−cont] feature value to [+cont] – is implemented in the phrasal phonology through the high ranking of a positional markedness constraint forbidding quasi-intervocalic [k]. It is reasonable to assume here that the constraint responsible for eliminating candidate (9b-i) from the running is a dialect-specific one: this is the case since GT itself is a dialect-specific phenomenon. Thus, the development from grammar 1 to grammar 2 involves speakers constructing and ranking a new positional markedness constraint that is phonetically grounded on the gradient spirantisation pattern that had previously been governed by implementational constraints in the phonetic module.[27]

Furthermore, it is worth pointing out that stabilisation of GT in grammar 2 need not mean that spirantisation comes under the *exclusive* control of the phrasal phonology. One possibility, illustrated in Figure 4.13 by the phonetic representations in grey, is that categorical spirantisation may exist in the grammar alongside a gradient version of the same process. This type of diachronic *scattering* of a process as it moves from one grammatical module to another over time is a specific prediction of the life cycle model (see Bermúdez-Otero & Trousdale 2012: 696).[28] One of the advantages of this for the analysis of a variable phenomenon like GT relates to the issue of stylistic variation. As shown in (9b), high ranking of the new markedness constraint that penalises intervocalic [k] relative to the faithfulness constraint that demands preservation of input manner-of-articulation features generates forms with intervocalic fricatives on the surface. However, the opposite ranking of these constraints shown in (9c) yields the same pattern as the conservative grammar in (9a). This opens up a range of options for speakers of the more innovative dialect. If we assume the ranking in (9b) is present in the co-phonology that is selected in casual speaking registers, then the alternative co-phonology containing the ranking in (9c) is selected in formal and in semi-formal registers. In the formal style, phonetic implementation of the winning candidate (9c-i) involves mapping in the intervocalic stop onto an Auditory Form and then an Articulatory Form that will result in the production of a fully occluded stop. By contrast, in more semi-formal speech, a different implementational pathway may be followed: as illustrated by Figure 4.14 below, implementation of (9c-i) can still generate a pattern of gradient

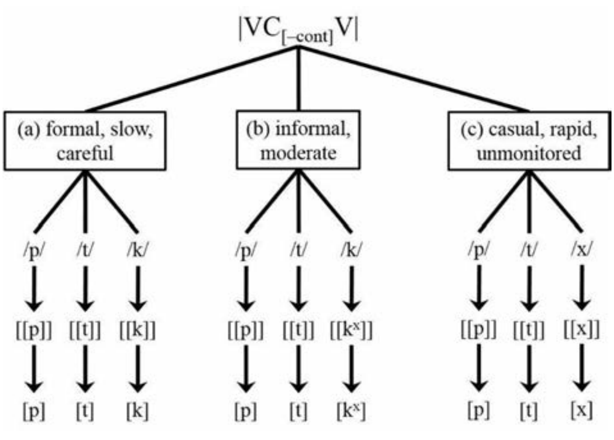


Figure 4.14 Stylistic variation of spirantisation in the innovative grammar. In the most casual speaking register (c), spirantisation is categorical and under the control of the phonological module. In the formal register (a), spirantisation is blocked both by the style-specific ranking of markedness and faithfulness in the categorical phonology and by implementational constraints in the phonetic module. In a more informal register (b), the constraint ranking in the phonological module prevents categorical spirantisation; however, assuming a stabilisation scenario in which spirantisation becomes synchronically scattered between the phonological and phonetic modules, some amount of controlled, gradient spirantisation may still arise in phonetic implementation through style-specific ranking of auditory and articulatory constraints

spirantisation if the style-specific ranking of auditory and articulatory constraints remains intact following stabilisation of spirantisation. This means that, if the spirantisation is scattered between the categorical phonology and the phonetic implementational module (as in grammar 2 in Figure 4.13), then grammar 2 is capable of generating *both* categorical and gradient patterns of spirantisation synchronically.

Scattering is also important for understanding rule generalisation. Once /k/-spirantisation has become a stable phrase-level process that is implemented by the casual register co-grammar, how then do /t/ and /p/ become targets for spirantisation? Crucially, the life cycle makes very restrictive claims about how this sort of expansion may operate which centre on the notion of *unidirectionality* in phonological change. As noted in section 4.2.1, the synchronic grammar assumed in modular frameworks of grammar function *bidirectionally*: in speech production, linguistic structures are fed forward from one module of grammar to another, whereas in speech comprehension, the process is reversed (i.e. comprehension involves 'feed-back' computation). In contrast to this, the life cycle assumes that phonological change typically only occurs in a feed-back direction (see Ramsammy 2015: section 2; Bermúdez-Otero & Trousdale 2012). That is, innovations first arise in phonetics and then may stabilise and become dependent on morpho-syntactic structure over time by domain narrowing. Other things being equal, this direction of change cannot occur in reverse order: the life cycle makes the strong claim that categorical phonological processes do not 'destabilise' and become gradient over time; and likewise, controlled, gradient phonetic processes are not predicted to 'de-phonologise' over time (see Bermúdez-Otero & Trousdale 2012 for discussion of parallels between the life cycle and processes of grammaticalisation and degrammaticalisation).

In strict adherence to the claims of the life cycle, the assumption of unidirectionality of change means that a categorical phonological process that spirantises intervocalic /k/ in casual speech styles cannot simply expand its structural description to target /t/ and /p/. On the contrary, /t/-spirantisation, and subsequently /p/-spirantisation, must go through the same phases of phonologisation and stabilisation as the original innovative process that gave rise to spirantisation of /k/ historically. This pathway of change is illustrated in Figure 4.15 below.

At timepoint 3, the grammar of Tuscan Italian is identical to that shown in Figure 4.14: /k/ spirantises categorically to [x] in casual speech but only gradiently in semi-formal style; and /t/ and /p/ never spirantise. By timepoint 4, however, a diachronic re-ranking of implementational constraints means that spirantisation has come to affect /t/: observe that whilst the categorical phonology preserves non-continuant realisations of /t/ in the output, gradient spirantisation can now apply in phonetic implementation. Here, the degree of spirantisation is somewhat more extreme in casual registers (i.e. the realisation may approach [θ]), whereas

	Output of categorical phonology				Output of phonetic implementation		
	formal	*infrm.*	*casual*		*formal*	*infrm.*	*casual*
Timepoint 3	/p, t, k/	/p, t, k/	/p, t, **x**/	⇛	[p, t, k]	[p, t, $\mathbf{k}^{\mathrm{x}}$]	[p, t, **x**]
Timepoint 4	/p, t, k/	/p, t, k/	/p, t, **x**/	⇛	[p, t, k]	[p, $\mathbf{t}^{\theta}$, $\mathbf{k}^{\mathrm{x}}$]	[p, **θ**, **x**]
Timepoint 5	/p, t, k/	/p, t, k/	/p, t, **x**/	⇛	[p, t, k]	[$\mathbf{p}^{\phi}$, $\mathbf{t}^{\theta}$, $\mathbf{k}^{\mathrm{x}}$]	[**φ**, **θ**, **x**]
Synchronic Tuscan	/p, t, k/	/**φ**, **θ**, **x**/	/**φ**, **θ**, **x**/	⇛	[p, t, k]	[**φ**, **θ**, **x**]	[**φ̞**, **θ̞**, **x̞**]

Figure 4.15 Diachronic generalisation of spirantisation

in semi-formal speech, frication is less extreme (i.e. a fricated stop or affricate, represented here as [t^{θ}]). This trend continues over time: whilst only /k/ and /t/ display spirantisation at timepoint 4, a further re-ranking of implementational constraints means that, by timepoint 5, /p/ follows the same pattern as /t/. By this stage, therefore, spirantisation of /t/ and /p/ has undergone phonologisation: speakers now control whether gradient lenition of these stops applies more or less extremely on a style-specific basis.

The life cycle then predicts that this gradient process will eventually stabilise into a fully fledged phonological rule as successive generations of speakers again reinterpret the gradient pattern as a categorical one. This change in the status of the spirantisation process gives rise to a grammar in which *all* underlying [−cont] obstruents map to [+cont] correspondents intervocalically in the output of the phonology in non-formal speech styles. As shown in (10) below, this can be captured through modification of the superordinate constraint militating against intervocalic [k] at timepoint 3 to one which prevents faithful output mappings of /p, t, k/.

(10)

a. Phrasal co-phonology for casual speech			
PL input: /la#kapɾa/ 'the goat'	$*[VC_{[-cont]}\ \{G_0,L_0\}V]$	IDENT-MoA	$*C_{[-cont]}$
i. [lakapɾa]	*!*		**
ii. [laxaɸɾa] ☜		**	

b. Phrasal co-phonology for formal speech			
PL input: /la#kapɾa/ 'the goat'	IDENT-MoA	$*[VC_{[-cont]}\ \{G_0,L_0\}V]$	$*C_{[-cont]}$
i. [lakapɾa] ☜		**	**
ii. [laxaɸɾa]	*!*		

Observe in (10b) that even in the most advanced dialects, the grammar still enforces the conservative pattern of faithful input→output mapping of underlying stops in the most formal speaking styles through the low ranking of the innovative spirantisation constraint in the formal co-phonological ranking. Nevertheless, the non-formal co-phonology shown in (10a) represents a grammar capable of generating the synchronic patterns of categorical spirantisation listed in Figure 4.15.

If the generalisation of spirantisation worked in this way, we should reasonably expect to observe a progressive increase in the use of spirantised forms over time. Indeed, currently available data have enabled us to propose a three-way theoretical reconstruction of the historical development of GT. We have hypothesised, firstly, that spirantisation has developed *grammatically* from an epiphenomenal by-product of fluent speech into a controlled gradience phonetic process, and from there into a fully fledged discrete phonological alternation. Secondly, spirantisation has extended *geographically* in parallel with the phonological innovations. In theoretical terms, the interpretation of the areal patterns is that /p,t,k/-spirantisation has undergone phonologisation and stabilisation in the most advanced dialects spoken in and around the city of Florence. However, the patterns observed in more peripheral areas are less advanced: dialects spoken in these regions maintain a more restricted use of spirantised forms that would have been typical in a previous historical stage of the synchronically most advanced Florentine dialects. Thirdly, the grammatical and geographical developments parallel the evolution of spirantisation in terms of progressive relaxation of its *stylistic limitations*. Whereas lenition was once restricted to the most casual and unmonitored speech styles, the ascension of the spirantisation process through the modular grammar and its areal spread go hand-in-hand with a relaxation of constraints militating against lenited forms. Thus, in the advanced Florentine dialect, spirantisation always occurs to some extent in casual and informal speech: synchronically, it is only the formal co-grammar that preserves the historical pattern of favouring fully occluded stops over spirantised variants.

4.7 Conclusion

This chapter has explored some important theoretical issues surrounding the phonology–phonetics interface from the perspective of constraint-based grammar. Our case study, *Gorgia Toscana*, is a well-known phenomenon that presents particular difficulties for analysis in constraint-based phonology. Currently available data show that GT is a highly variable phenomenon that displays both discrete phonological and continuous phonetic characteristics. The occurrence of variable stop spirantisation has been shown to correlate with dialectal factors, stylistic factors and sociolinguistic factors in addition to specifically phonological factors (e.g. place of articulation, phonotactic context etc.). The aim of this chapter has therefore been to construct an analysis of (some of) the patterns of variation that GT presents using the toolkit provided by two contemporary theories of phonology that have grown out of the optimality-theoretic research tradition.

Bidirectional Phonology and Phonetics is a framework that defines specific links between categorical phonological computation and larger cognitive processes involved in speech production and speech comprehension. The explicit claims that BiPhon makes about how phonological representations are interpreted by the phonetic implementational module and how phonological representations are transformed into auditory and articulatory objects through constraint-evaluation mechanisms make it an ideal framework for modelling a phenomenon like GT which simultaneously displays gradient and categorical characteristics. Furthermore, the life cycle of phonological processes allows for questions relating to how phonology and phonetics interact diachronically to be addressed. This model is based on a clearly defined set of claims about the architecture of grammar which, in turn, allow for hypotheses to be made about the interleaving of morphological and phonological structure, and about how the morpho-phonology, as a module of grammar, interfaces with phonetic implementation. This opens up a range of possibilities, many of which have a direct relevance for analysing phenomena like GT. For example, the life cycle allows the analyst to respond to theoretical questions such as where controlled gradient patterns come from, and how gradient patterns become categorical over the course of time. If the goal of phonology is to understand the aspects of sound patterns that are under the cognitive control of speakers, then explanation of gradient patterns is no less important than explanation of discrete allophonic patterns. Thus, phonologists fully committed to providing answers to questions posed both by synchronic language data and phonological patterns arising from historical change need theories that are equipped to deal with gradient and categorical patterns. The synthesis of the fundamental claims of the life cycle with the explicit constraint-based interface mechanisms of BiPhon thus provide one way of addressing questions about variability, gradience and phonological change that other theories of phonology – both within the realm of constraint-based theory and beyond – have had to steer clear of, or have otherwise ignored.

There remains, nevertheless, much work to do if we are to arrive at a theory of the phonology–phonetics interface capable of handling all the complexities of language data such as those presented by GT. Nothing has been said here, for example, about how variation in the use of spirantisation may relate, to some extent, to word frequency; and likewise, whilst there is evidence for sociolinguistic variation in the use of GT, this is not something that typically figures in phonological accounts (including the analyses presented in this chapter). Arriving at a fuller understanding of GT – or, indeed, of any phonological process – therefore entails, on the one hand, developing theoretical models that can cope with noisy data. On the other hand, there is a crucial need for sophisticated empirical research in order to provide the most complete and detailed descriptions of phonological phenomena. With regard to GT, the currently available data supply a tantalising but very incomplete picture about the phonetic processes involved in spirantisation; and there is little in-depth, large-scale research on frequency patterns and sociolinguistic variation. The dearth of data is a significant limitation on the progression of theoretical work: good phonetically sophisticated descriptions of language-specific phenomena are of inestimable value for phonological theory. Further development of theories of the phonology–phonetics interface will therefore rely heavily on what the results of carefully executed phonetic research can tell us about the complexities of speech production and perception. Uncovering and recognising the ways in which existing theoretical frameworks are ill equipped to handle specific empirical problems should therefore push us into thinking how those frameworks might be modified and optimised to provide the most complete and most coherent explanations of sound patterns in language.

The question of how phonology and phonetics interact synchronically and diachronically thus remains a central and critical question for theoretical phonology. As stated at the outset of this chapter, approaching the sorts of questions posed by GT data from a constraint-based perspective is but one of the options available to phonologists in the 21st century. Nevertheless, developing new accounts of complex interface phenomena like GT constitutes an important direction for future research in theoretical phonology, both within the optimality-theoretic tradition and otherwise.

Notes

I am grateful to Silke Hamann and Ricardo Bermúdez-Otero for providing useful comments on a previous version of this paper. Any remaining errors are my own.

1 For overviews on theories of learnability in OT, see Tesar & Smolensky (2000); Boersma & Hayes (2001); Boersma, Escudero & Hayes (2003); Turton (2012). See also Iosad (this volume: section 2.4) and Krämer (this volume: section 3.1) for some discussion of learnability.

2 I assume here that voicing of [z] in the [-ozo] suffix arises through a phonological process that changes the laryngeal specification of underlying /s/. See Iosad (this volume: section 1.3) for a similar case of intervocalic /s/ voicing that is unrelated to the spirantisation patterns observed in Tuscan Italian to be discussed later in this chapter. See also Krämer (2009: 207–219) for a discussion of the operation of intervocalic /s/-voicing in Italian.

3 This brief analysis of truncation is based on consideration of other forms like *libraccio* 'dreadful book', *librario* 'of books' and *libreria* 'bookshop' (/líbɾo-átʃːjo/, /líbɾo-áɾjo/ and /líbɾo-ería/, respectively). In each case, stem-final /o/ (i.e. V_1) is deleted in favour of preserving the suffix-initial vowel (i.e. V_2). Moreover, since /líbɾo-ería/ displays a sequence of two unstressed vowels spanning a morpheme boundary, it seems unlikely that truncation is driven by stressed vowel dominance (which may be suggested by forms with stressed suffix-initial vowels like /líbɾo-étːo/, /líbɾo-átʃːjo/ and /líbɾo-áɾjo/: cf. Peperkamp 1995: 210). Nevertheless, some support that stress placement plays a role in vowel deletion comes from compounds like *porta abiti* 'clothes rail' in which truncation is blocked to avoid stress clash ([pòɾtaábiti] is well formed, whereas *[pòɾtábiti] is not; cf. *porta ombrelli* [pòɾtombɾélːi] 'umbrella stand'). See Krämer (2009: section 6.4) for further discussion.

4 I assume that the /-oso/ suffix which derives *pericoloso* from *pericolo* ‘danger’ attaches at the stem level. Some evidence for this comes from the fact that /-oso/ can occur before other suffixes like superlative /-is:imo/, hence *pericolosissimo* ‘most dangerous’. Furthermore, that this suffix contains an initial /o/ is clear from forms like /fama-oso/→[famozo] ‘famous’. As in the case of words like /libɾo-et:o/, V_1 in the /V_1V_2/ sequence created by concatenation of stem and suffix in /peɾikolo-oso/ deletes to prevent ill-formed hiatus, i.e. *[peɾikoloozo].

5 Kirchner (2004) makes use of the term ‘quasi-intervocalic’ to refer to sequences like /-ibɾo/ in which the /b/ is a target for spirantisation despite the presence of a following sonorant consonant in the phonological string. I assume here that the positional markedness constraint that triggers spirantisation assigns violations to stops in quasi-intervocalic environments. For clarity, it is worth mentioning that Iosad (this volume: section 2.1.3) employs a slightly different constraint that penalises *post-vocalic* stops (namely $^*VC_{[-son,\ -cont]}$).

6 Marotta (2008: 239) notes that consonants other than stops may also be weakened in Tuscan dialects (e.g. the affricates /ʧ/ and /ʤ/). I focus exclusively on the voiceless stops throughout this chapter. In so doing, the goal is not to provide a fully detailed account of everything of phonological relevance for GT, but rather to illustrate how the variable aspects of stop spirantisation can be accounted for in constraint-based frameworks.

7 Specifically, in the *Atlante linguistico-etnografico dell’Italia e della Svizzera meridionale* (*AIS*, henceforth).

8 Note that this diachronic scenario also mirrors the rates of synchronic spirantisation recorded in Gianelli & Savoia’s (1978) study: i.e. /k/ > /t/ > /p/.

9 See Marotta (2008: 245ff.) for some details on the differences in the application of GT in Florence (the locus of the innovations) and Pisa (a more peripheral location along the diffusion trajectory shown in Figure 4.4).

10 Gianelli & Savoia (1978: 35–39) also mention debuccalised realisations of /t/ and /p/: e.g. /t/ in *bevuto* ‘drunk’ may occur as [x], [h] or [∅], and likewise /p/ in *lo prende* ‘s/he takes it’ can occur as [ɸ] or [h].

11 Villafaña-Dalcher (2008), following Marotta (2001), makes reference to ‘semi-fricative’ realisations rather than affricates. The latter term is employed throughout this chapter for the sake of consistency with other work.

12 Note that Villafaña-Dalcher’s data do not exactly mirror the patterns in the *AIS*. Whilst /k/ displays the most robust levels of spirantisation, /p/ shows a tendency to spirantise somewhat more frequently than /t/. However, as Marotta (2008: 250, footnote 13) notes, this result is inconsistent with other studies on GT. One explanation for this may be that Villafaña-Dalcher’s study is based on the speech of only six Florentine speakers: a broader study surveying the use of spirantised stops from a greater number of speakers across locations in Tuscany may yield a picture more consistent with other studies.

13 Articulatory targets are represented by horizontal ‘windows’ defined by pairs of solid black lines in Figure 4.7. Observe that stops like /k/ have a very narrow window which represents the target for full linguo-velar occlusion.

14 See also Solé (2007) and references cited therein for relevant discussion of the distinction between automatic vs cognitively-controlled phonetic processes.

15 Observe here that BiPhon makes use of specific bracketing conventions: URs are enclosed in |vertical bars|; phonological surface forms, i.e. output structures generated by the categorical phonology, are enclosed in /slanted lines/; Auditory Forms appear in [[double square brackets]]; and Articulatory Forms appear in [single square brackets]. I adhere to these conventions in this chapter when discussing this framework specifically. However, I use more standard notation, i.e. /input/→[output], for the sake of simplicity where the intended meaning is unambiguous.

16 In addition, dialect-specific and even idiosyncratic speaker-specific factors may influence the ranking of cue constraints.

17 See Halle, Hughes & Radley (1957) on the expected acoustic properties of the burst in stops at different places of articulation.

18 In Figure 4.10 and in following tableaux, I employ > as a shorthand for the approach pinch and < to represent the release pinch.

19 F1 and F2 values used here are idealisations for Standard Italian based on measurements published in Ferrero (1972).

20 Note that the cue constraints in (6) have been abbreviated for convenience. The key idea behind these constraints is that they evaluate auditory candidates on the basis of closest approximation to all perceptually relevant acoustic landmarks (i.e. including frequencies of higher formants in vowels; and including burst duration, burst amplitude, closure-silence duration, VOT, etc. in stops).

21 See Gafos (2002: 270) for the relevant statement: "Principles or constraints *in the grammar* refer to temporal relations between gestures" (emphasis in original).

22 It is assumed here that candidate (7c) also respects other low ranked constraints which have been omitted from the tableau. For example, the first vowel in [œkæ] would incur no violation of a constraint (e.g. [[ɔ]]→TB:PAL) permitting the auditory quality of [[ɔ]] to be approximated by a vowel articulation involving a more anterior oral constriction as in [œ].

23 It should be noted here that this is also dependent on the correct aerodynamic conditions being created by articulatory configurations.

24 For clarity, cue constraints responsible for evaluating mappings of vowels in the SR to auditory candidates have been omitted from (8); cf. (6).

25 Note that I use the term co-phonology in the sense of alternative rankings of constraints in a particular module of grammar that are activated synchronically depending upon speaking style. In this case, the proposal is that phonetic implementational constraints that are ranked into style-specific hierarchies such that *co-grammar* may be a more appropriate way of conceptualising this idea. It should be highlighted that this use of the term co-phonology differs fundamentally from its use in analyses of morpho-phonological patterns: cf. Inkelas & Zoll (2007) and references cited therein.

26 For further discussion of the role of *input restructuring* in the life cycle, see Ramsammy (2015: section 2) and Bermúdez-Otero (2006, 2007: section 21.3.2).

27 This analysis therefore takes on board the hypothesis that the constraints in CON are not universal, but rather that they are created on a language-specific basis by the learner in acquisition: see Iosad (this volume: section 2.2.1) for comments. The advantage of this approach is that it avoids furnishing the grammar of non-spirantising dialects like Standard Italian with superfluous constraints that never trigger any phonological alternations.

28 In this sense, the apparent 'duplication' of a process between two grammatical modules is not viewed as problematic in the life cycle (cf. Iosad, this volume: section 2.2.4). As the gradient and categorical versions of spirantisation affect representations in different ways, the possibility of scattering is something that can be exploited by the analyst to account for phonological phenomena that simultaneously display continuous and discrete characteristics synchronically.

References

Anderson, Stephen R. 1981. Why phonology isn't "natural". *Linguistic Inquiry* 12/4: 493–539.

Bermúdez-Otero, Ricardo. 2006. Phonological change in Optimality Theory. In Keith Brown (ed) *Encyclopaedia of language and linguistics*, volume 9 (2nd ed.). Oxford: Elsevier, 497–505.

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In Paul de Lacy (ed) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 497–517.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds) *The Blackwell companion to phonology*, volume 4. Malden, Massachusetts: Wiley-Blackwell, 2019–2048.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed) *The morphology and phonology of exponence*. Oxford: Oxford University Press, 8–83.

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus* 25/1: 3–103.

Bermúdez-Otero, Ricardo. 2014. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds) *The Oxford handbook of historical phonology*. Oxford: Oxford University Press, 374–399.

Bermúdez-Otero, Ricardo & April McMahon. 2006. English phonology and morphology. In Bas Aarts & April McMahon (eds) *The handbook of English linguistics*. Oxford: Blackwell, 382–410.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds) *The Oxford handbook of the history of English*. Oxford: Oxford University Press, 691–720.

Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989–2054.

Boersma, Paul. 2009. Cue constraints and their interactions in phonological perception and production. In Paul Boersma & Silke Hamann (eds) *Phonology in perception*. Berlin: Mouton de Gruyter, 55–110.

Boersma, Paul, Paola Escudero & Rachel Hayes. 2003. Learning abstract phonology from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. *Proceedings of the 15th International Congress of Phonetic Sciences*: 1013–1016.

Boersma, Paul & Silke Hamann. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32/1: 45–86.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3: 219–252.

Browman, Catherine P. & Louis M. Goldstein. 1990. Tiers in articulatory phonology with some implications for casual speech. In John Kingston & Mary Beckman (eds) *Papers in laboratory phonology I: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press, 341–376.

Browman, Catherine P. & Louis M. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49/3–4: 155–180.

Bradley, Travis. 2006. Spanish complex onsets and the phonetics–phonology interface. In Fermando Martínez-Gil & Sonia Colina (eds) *Optimality-theoretic studies in Spanish phonology*. Amsterdam: John Benjamins, 15–38.

Buizza, Emanuela & Leendert Plug. 2012. Lenition, fortition and the status of plosive affrication: The case of spontaneous RP English /t/. *Phonology* 29: 1–38.

Cavirani, Edoardo. 2015. *Modelling phonologization: Vowel reduction and epenthesis in Lunigiana dialects*. PhD thesis, Leiden University.

Cohn, Abigail. 1993. Nasalisation in English: Phonology or phonetics? *Phonology* 10: 43–81.

Collie, Sarah. 2007. *English stress preservation and Stratal Optimality Theory*. PhD thesis, University of Edinburgh.

Ferrero, Franco. E. 1972. Caratteristiche acustiche dei fonemi vocalici italiani. *Parole e metodi: bollettino dell'Atlante linguistico italiano* 3/1: 9–32.

Gafos, Adamantios I. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20: 269–337.

Gianelli, Luciano & Leonardo M. Savoia. 1978. L'indebolimento consonantico in Toscana: I. *Rivista italiana di dialettologia* 2: 23–58.

Gianelli, Luciano & Leonardo M. Savoia. 1979–80. L'indebolimento consonantico in Toscana: II. *Rivista italiana di dialettologia* 3/4: 38–101.

Hall, Robert A. 1949. A note on the "Gorgia Toscana". *Italica* 26/1: 64–71.

Halle, M., G. W. Hughes & J.-P. A. Radley. 1957. Acoustic properties of stop consonants. *Journal of the Acoustical Society of America* 29/1: 107–116.

Hamann, Silke. 2010. Phonetics–phonology interface. In Nancy C. Kula, Bert Botma & Kuniya Nasukawa (eds) *Continuum companion to phonology*. London: Continuum International Publishing Group, 202–224.

Honeybone, Patrick. 2012. Lenition in English. In Terttu Nevalainen & Elizabeth Closs Traugott (eds) *The Oxford handbook of the history of English*. Oxford: Oxford University Press, 773–787.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45/1: 133–171.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Keating, Patricia. A. 1990. The window model of coarticulation: Articulatory evidence. In John Kingston & Mary E. Beckman (eds) *Papers in laboratory phonology I: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press, 451–470.

Kingston, John. 2011. Tonogenesis. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds) *The Blackwell companion to phonology*, volume 4. Malden, Massachusetts: Wiley-Blackwell, 2304–2333.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In Yang, I. S. (ed) *Linguistics in the morning calm*. Seoul: Hanshin, 3–91.

Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2: 85–138.

Kirchner, Robert. 2004. Consonant lenition. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds) *Phonetically-based phonology*. Cambridge: Cambridge University Press, 313–345.

Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.

Marotta, Giovanna. 2001. Non solo spiranti: la 'gorgia toscana' nel parlato di Pisa. *L'Italia dialettale* 62: 27–60.

Marotta, Giovanna. 2008. Lenition in Tuscan Italian (Gorgia Toscana). In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds) *Lenition and fortition*. Berlin: De Gruyter Mouton, 235–272.

McCarthy, John. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

Mohanan, K. P. 1986. *The theory of Lexical Phonology*. Dordrecht: Reidel.

Montemagni, Simonetta, Martijn Wieling, Bob de Jonge & John Nerbonne. 2013. Synchronic patterns of Tuscan phonetic variation and diachronic change: Evidence from a dialectometric study. *Literary and Linguistic Computing* 28/1: 157–172.

Ohala, John. 1993. The phonetics of sound change. In Charles Jones (ed) *Historical linguistics: Problems and perspectives*. London: Longman, 237–278.

Peperkamp, Sharon. 1995. Prosodic constraints in the derivational morphology of Italian. *Yearbook of Morphology* 1994: 207–244.

Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds) *Laboratory phonology 7*. Berlin: Mouton de Gruyter, 101–139.

Proctor, Michael I. 2009. *Gestural characterisation of a phonological class: The liquids*. PhD thesis, Yale University.

Ramsammy, Michael. 2015. The life cycle of phonological processes: Accounting for dialectal microtypologies. *Language and Linguistics Compass* 9/1: 33–54.

Scobbie, James M. 2007. Interface and overlap in phonetics and phonology. In Gillian Ramchand & Charles Reiss (eds) *The Oxford handbook of linguistic interfaces*. Oxford: Oxford University Press, 17–52.

Simpson, Adrian P. 2001. Does articulatory reduction miss more patterns than it accounts for? *Journal of the International Phonetic Association* 31/1: 29–39.

Solé, Maria-Josep. 2007. Controlled and mechanical properties in speech: A review of the literature. In Maria-Josep Solé, Patrice Beddor & Manjari Ohala (eds) *Experimental approaches to phonology*. Oxford: Oxford University Press, 302–321.

Sorianello, Patrizia. 2002. I suoni fricativi dell'italiano fiorentino. In S. Calamai & C. Finocchiaro (eds) *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa* 3: 26–39.

Tesar, Bruce & Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge: Cambridge University Press.

Turton, Danielle. 2012. The darkening of English /l/: A stochastic Stratal OT analysis. Unpublished MS, University of Manchester (http://ling.auf.net/lingBuzz/001524).

Vennemann, Theo. 1978. Phonetic analogy and conceptual analogy. In Philip Baldi & Ronald N. Werth (eds) *Readings in historical phonology: Chapters in the theory of sound change*. Philadelphia: Pennsylvania State University Press, 258–274.

Villafaña-Dalcher, Christina. 2008. Consonant weakening in Florentine Italian: A crossdisciplinary approach to gradient and variable sound change. *Language Variation and Change* 20: 275–316.

5
Stratal Phonology

Ricardo Bermúdez-Otero

5.1 Introduction

Stratal Phonology is a theory of how phonology interacts with other components of grammar. Its basic principles are simple: phonology applies cyclically over domains defined by certain constituents in the morphosyntactic structure of linguistic expressions, and domains associated with constituents of different rank – stems, words, and utterances – obey different phonological generalizations. In current versions of the framework (e.g. Kiparsky 2000, 2015b; Bermúdez-Otero 2010), these hypotheses are combined with constraint-based models of phonological computation, like Optimality Theory ('OT': Prince & Smolensky 1993; Iosad, this volume; Krämer, this volume).

The hypotheses of cyclicity and stratification are laid out in §5.2. The assumption of cyclic application predates the rise of Stratal Phonology and provides some common ground with several other approaches to phonology's upper interfaces: notably, Cophonology Theory (Orgun 1996; Inkelas 1998, 2012) and various phonological applications of Chomsky's (2001) Phase Theory (e.g. Embick 2014; D'Alessandro & Scheer 2015; Newell, this volume). Stratal Phonology differs from these in positing relatively fewer cycles. The theory also diverges in important ways from its most immediate precursor: rule-based Lexical Phonology (Pesetsky 1979; Kiparsky 1982a, 1982b; Mohanan 1982). First, Stratal Phonology rejects the claims of Strict Cyclicity and Structure Preservation, which sought to constrain the application of rewrite rules at the stem level. Secondly, Lexical Phonology simply stipulated a number of important generalizations about cyclic domain structures, such as the fact that roots do not define cyclic domains and that stem-level domains are recursive; in contrast, recent work in Stratal Phonology seeks to derive these observations from independent facts. Throughout §5.2 I emphasize the major empirical predictions of Stratal Phonology, which include Cyclic Containment, the Russian Doll Theorem, and Chung's Generalization.

Cyclic Containment[1] holds that, in cases of morphosyntactically induced phonological opacity, a linguistic expression inherits its opaque phonological properties from a constituent that defines an immediate cyclic subdomain. In recent years, the proponents of output–output correspondence (henceforth 'OO-correspondence') have adduced a number of putative counterexamples to this prediction. The theory of OO-correspondence asserts, instead, that the phonological computation may directly refer to a surface base that does not match a constituent of the opaque expression (e.g. Kenstowicz 1996; Burzio 1996; Steriade 1999). In §5.3 I address this debate, highlighting the divergent empirical predictions of the cycle and OO-correspondence. As a test case, I pay particular attention to Steriade's (1999: §2–§3) discussion of English affixes with dual-level behaviour, notably *-able*. Steriade's analysis uncovers genuine and previously underappreciated empirical facts, in which the paradigmatic relationships highlighted by transderivational approaches do play a key role. I shall argue, however, that this role should be regarded as taking effect during lexical and morphological acquisition, rather than in the phonological derivation.[2] This account, when implemented within the framework of Stratal Phonology, predicts certain empirical observations that are not captured by OO-correspondence, such as the fact that stress-affecting instances of *-able* suffixation like *re.mé.dĭ.a.ble* (cf. *rémedy*) exhibit the same metrical pattern as forms based on bound roots (e.g. *in.dó.mĭ.ta.ble*). From this and other considerations I conclude that cyclicity retains an empirical advantage over OO-correspondence.

Finally, §5.4 asks whether Stratal Phonology permits a graceful integration with other components of grammar, particularly morphology. The theory would be in trouble if it made demonstrably false assumptions about morphology, or if it crucially relied on excessively powerful exponence mechanisms that robbed morphological theory of its empirical content. In this connection, I show, first, that the serial precedence of the stem-level phonology over the word-level phonology does not depend on level ordering, understood as the requirement that all stem-level affixes should occur inside all word-level affixes (cf. Kiparsky 1982a: 131ff.). Similarly, Stratal Phonology need not resort to rebracketing operations to deal with so-called 'bracketing paradoxes' (cf. Kiparsky 1983: §5). More fundamentally, however, Stratal Phonology does presuppose that it is possible to demarcate morphosyntax from phonology, for it claims that the morphosyntactic operations in a language can be sorted into a small number of classes (called 'levels' or 'strata') according to the phonological processes for which they define cyclic domains. In opposition to other cyclic frameworks like Amorphous Morphology (Anderson 1992) and Cophonology Theory, I suggest that the best way of delimiting the roles of morphology and phonology in exponence is by adopting a strictly modular stance, in which morphology can select and insert morphs, but cannot alter their phonological content. This, in turn, favours approaches to apparently nonconcatenative exponence along the lines of Generalized Nonlinear Affixation (Bermúdez-Otero 2012: 53).

The implications for morphology reviewed in §5.4 do not exhaust the predictions of Stratal Phonology. The theory has important consequences for many other domains of enquiry. A selection of references is included in the further reading suggestions at the end of this chapter.

5.2 Basic principles of Stratal Phonology

5.2.1 Demarcating the framework

The main ideas behind Stratal Phonology have a long and complex intellectual history. According to Kiparsky (1983: 3), the distinction between stem-level and word-level affixation can be traced back to Pāṇini by way of Bloomfield's (1933: 209ff, 1939: §6–§9) 'primary' and 'secondary' affixes. Similarly, Booij (1997: 264) observes that the distinction between 'lexical' and 'postlexical' phonology was already codified in the Praguian terms *phonologie du mot* and *phonologie de la phrase* (CLP 1931: 321; Jakobson 1931: 165). The phonological cycle, in turn, is as old as generative phonology itself (Chomsky, Halle &

Lukoff 1956: 75). The closest ancestor of current stratal work is to be found in rule-based Lexical Phonology and Morphology. As noted in §5.1, however, research in Lexical Phonology paid a great deal of attention to principles like Strict Cyclicity and Structure Preservation, which governed rule application at the stem level; these hypotheses have since been abandoned (Bermúdez-Otero 2013b). Nonetheless, rule-based stratal theories descending from Halle & Vergnaud (1987) remain in use, particularly in work associated with Distributed Morphology ('DM': Halle & Marantz 1993, 1994): see e.g. Embick (2014) and Newell (this volume).

In this chapter, therefore, the term 'Stratal Phonology' is strictly reserved for work that combines a stratal phonological architecture with contemporary constraint-based parallelist approaches to phonological mappings. This includes not only Stratal OT, but also frameworks where phonological generalizations are expressed by means of Harmonic Grammar (Pater 2009) or Maximum Entropy ('MaxEnt') modelling (Hayes & Wilson 2008): see e.g. Nazarov & Pater (2017) for a stratal MaxEnt study in the acquisition of opacity. In contrast, the requirement of parallelism excludes a potential combination of stratification either with Harmonic Serialism (McCarthy 2010) or with OT with Candidate Chains (McCarthy 2007). This exclusion is motivated by the fact that the hypothesis of cyclic derivation, which is absolutely central to Stratal Phonology, loses much of its empirical content in frameworks that adopt a serialist approach to phonological mappings: see Bermúdez-Otero (2013a: 90–1) for an example. Similarly, current constraint-based parallelist theories are happily unable to express invalid claims like Strict Cyclicity or Structure Preservation, whereas statements like Chung's Generalization (§5.2.3.3) are derived as theorems.

5.2.2 The cycle

The key principles of Stratal Phonology are cyclicity and stratification.[3] For the purposes of elucidating the concept of the phonological cycle, let us think of morphology as establishing relationships of exponence between nodes in a syntactic structure and phonological pieces in an underlying representation (for specific proposals, see Bermúdez-Otero 2012: 46–8, 50–3). Phonology, in turn, maps the assembly of exponents built by the morphology onto a surface representation. In a cyclic framework, this mapping is in fact specified by a composite function. Following Kaye (1995: 302), we can conceive of phonological theory as defining a set of $\mathcal{P}$-functions mapping any given phonological input representation *i* onto a corresponding output *o*.[4] In OT, for example, the phonological function $\mathcal{P}_r(i) = o$ consists of an application of Gen followed by an application of Eval_r, where r is a ranking of the constraint set Con:

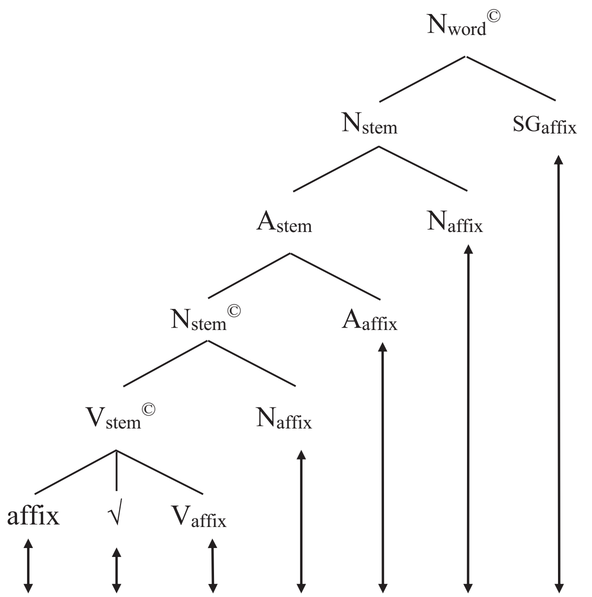
(1) $\mathcal{P}_r(i) = \textit{Eval}_r(\textit{Gen}(i)) = o$

Now certain nodes in the syntactic structure of a complex linguistic expression can be designated as 'cyclic', in the sense that the assembly of exponents associated with a cyclic node provides the argument for the application of a $\mathcal{P}$-function. Crucially, $\mathcal{P}$-functions triggered by higher cyclic nodes apply to the results of $\mathcal{P}$-functions triggered by lower cyclic nodes, so that the surface representation of the whole expression is obtained by function composition.

In (2), for example, morphology has associated the syntactic structure of the English singular noun *accommodationlessness* (2,a) with the underlying phonological representation in (2,b).[5] The relationships of exponence thus established are indicated by double-headed arrows. Cyclic nodes are highlighted with the superscript ©, and the corresponding cyclic domains in the underlying representation (2,b) are demarcated with square brackets [ ]. The subscripts $_{SL}$ and $_{WL}$ indicate the affiliation of affixes to the stem or word level (discussed in §5.2.3 below). Given all of the above, the surface representation is determined by the composite function shown in (2,c), with the phonological derivation proceeding as per (2,d).

(2)

a.



b. [[[a- commod -ate$_{SL}$] -ion$_{SL}$] -less$_{WL}$ -ness$_{WL}$ -∅ $_{WL}$]

c. $\mathcal{P}_{WL}$($\mathcal{P}_{SL}$($\mathcal{P}_{SL}$(a-, commod-, -ate), -ion), -less, -ness, -∅)

d.	$\mathcal{P}_{SL}$	1st cycle	accómmodàte
		2nd cycle	accòmmodátion
	$\mathcal{P}_{WL}$	3rd cycle	accòmmodátionlessness

The order of $\mathcal{P}$-function application is thus intrinsically determined by morphosyntactic constituency: the computation of the phonological form of the parts precedes and feeds the computation of the phonological form of the whole. Stratal Phonology derives a great deal of its empirical content from this simple notion. Notably, like all cyclic frameworks, stratal theories predict that morphosyntactically induced opacity is subject to Cyclic Containment:

(3) Cyclic Containment
In cases of morphosyntactically induced phonological opacity, a linguistic expression inherits its opaque phonological properties from a constituent defining an immediate cyclic subdomain.

The stress profile of the English word *accòmmodátionlessness*, for example, is doubly opaque: first, the word exhibits prefenestral primary stress (i.e. primary stress outside the final trisyllabic window); secondly, pretonic secondary stress fails to fall on the initial syllable (cf. monomorphemic items like *àbracadábra*). As shown in (2), this is because *accòmmodátionlessness* inherits the metrical contour of the noun stem *accòmmodátion-*, which defines an immediate cyclic subdomain: *accòmmodátion-* is a cyclic constituent, and there is no other cyclic node between *accòmmodátion-* and *accòmmodátionlessness*. In turn, *accòmmodátion-* inherits the foot-head on its second syllable from its own base, the verb stem *accómmodàte-*, which defines an immediate cyclic subdomain. This accounts for the ungrammaticality of **àccommodátion*. Observe that neither the adjective stem *accòmmodátionless-* nor the verb stem *accómmodàte-* defines an immediate cyclic subdomain within *accòmmodátionlessness*. The adjective stem is not a cyclic constituent (for reasons discussed in §5.2.3.2 below). The verb stem, though cyclic, is a remote or nonlocal base, rather than a proximate or local one; there is another cyclic node, the noun stem *accòmmodátion-*, closer to *accòmmodátionlessness*. The fact that *accòmmodátionlessness* is in an immediate cyclic relationship with *accòmmodátion-*, but not with *accómmodàte-*, explains the ungrammaticality of **accómmodàtionlessness*.

As we saw in §5.1, however, recent work on OO-correspondence has called Cyclic Containment into question: I examine and reject the putative counterevidence in §5.3 below.

Another entailment of cyclic theory is the Russian Doll Theorem:

(4) The Russian Doll Theorem (Bermúdez-Otero 2011: 2023)
Let there be the nested cyclic domains $[_{\gamma} \ldots [_{\beta} \ldots [_{\alpha} \ldots] \ldots] \ldots]$. If a phonological process p is opaque in β because its domain is α, then p is opaque in γ.

In the English derived adjective [[*long*]*ish*] [lɒŋɪʃ], for example, postnasal /g/-deletion overapplies before the initial vowel of the suffix *-ish* [ɪʃ] because its cyclic domain is the adjective stem *long-* [lɒŋ], which is contained within *longish*.[6] The Russian Doll Theorem correctly predicts that postnasal /g/-deletion will also overapply when a vowel follows across a word boundary, as in the phrase *long effect* [lɒŋɪfɛkt]: by simple transitivity, if opacity arises when a word contains a prevocalic token of the stem *long-*, it will also arise when a phrase contains a word that itself contains a prevocalic token of the stem *long-*. In noncyclic frameworks assuming OO-correspondence, in contrast, the Russian Doll Theorem is enforced by stipulation (Bermúdez-Otero 2011: 2043).

The extent to which the Russian Doll Theorem holds true has attracted surprisingly little comment. I know of only one putative counterexample, reported in Albright (2015).

5.2.3 Stratification

5.2.3.1 Key generalizations

The preceding characterization of the phonological cycle generalizes to a wide range of cyclic frameworks. Stratal Phonology, however, adds two other important claims. First, cyclic domain structure is sparse: relatively few morphosyntactic constituents trigger phonological cycles. Secondly, there are different $\mathcal{P}$-functions for cyclic nodes of different rank.

These ideas are developed in the theory of stratification. The latter makes crucial reference to the concepts of 'root', 'stem', and 'word'. For our current purposes, these may be defined as follows:

(5) a. A **root** (√) is a minimal acategorial lexical item.
b. A **stem** is a lexical item specified for syntactic category (N, V, A, etc.). In certain cases it is necessary to distinguish further between **derivational stems** and **inflectional stems**, where a derivational stem belongs to a syntactic category but must undergo some further morphosyntactic operation before it becomes inflectable.
c. A **word** is a syntactically autonomous lexical item bearing the full set of inflectional features required by its category.

In (2), for example, the verb stem *a-ccommod-ate-* derives from the root *commod-*, which by itself has no lexical category and is *a fortiori* uninflectable; cf. the adjective *commod-ious* and the noun *commod-ity*. The topmost node in the structure constitutes a word: namely, a noun covertly inflected for singular number (cf. *weakness-∅* ~ *weakness-es*). All the nodes intervening between the root and the word are stems: they are specified for lexical category but they are not inflectionally complete. For more extensive discussion of these concepts, see e.g. Kiparsky (2003a) and Bermúdez-Otero (2013a).

We can now formulate four key generalizations (cf. Bermúdez-Otero 2006: 283), to be slightly revised in (23):

(6) a. Roots do not define cyclic domains.
b. Some stems define cyclic domains for the stem-level phonology ($\mathcal{P}_{SL}$).
c. Words define cyclic domains for the word-level phonology ($\mathcal{P}_{WL}$).
d. Utterances define cyclic domains for the phrase-level phonology ($\mathcal{P}_{PL}$).

Most theories in the stratal tradition subscribe to these statements: for example, stem-level, word-level, and phrase-level phonological constraints as defined here correspond roughly to the cyclic, postcyclic, and postlexical phonological rules of Booij & Rubach (1987).[7]

Why the generalizations in (6) should hold is an interesting topic of research. Take, for example, the phonological inertness of roots (Kiparsky 1982a: 144–5, 1982b: 32–3; Inkelas 1989: §3.5). In an investigation of Spanish morphology, Bermúdez-Otero (2013a) adduces empirical evidence, both internal (the underlying distribution of theme vowels, locality effects in allomorph selection) and external (response latencies in lexical recognition), to show that stems are the minimal units of lexical storage and that all productive morphology is either stem-based or word-based. If these propositions are correct, they automatically explain why (6,a) holds true for Spanish (on Portuguese, see Matzenauer & Bisol 2016). Intriguingly, this approach leaves open the possibility that different stratification régimes may hold in languages that provide learners with better cues for root extraction, if such exist (Bermúdez-Otero 2013a: 53–4).

5.2.3.2 The nonrecursiveness of word-level and phrase-level domains

Cyclic domain structures conforming to the generalizations in (6) are relatively sparse. The phrase-level phonology, for instance, applies once across the board over the entire utterance; phrasal categories smaller than the utterance (e.g. DPs, IPs, etc.) do not trigger phonological cycles. This means that no phonological process applies in a cyclic domain smaller than the utterance but larger than the maximal grammatical word; domains appearing to satisfy this description are prosodic, not cyclic (for the difference, see Bermúdez-Otero & Luís 2009; Bermúdez-Otero 2011: §4).

Similarly, a word-level domain will rarely be found embedded within another word-level domain: this is because stems, when cyclic, undergo the stem-level phonology, whilst the grammatical word as defined in (5,c) rarely behaves as a recursive category.[8] In line with (6), therefore, only stem-level domains are ordinarily found nested within domains of the same type. We saw an example of this in (2). The noun stem *accommodation-* contains the verb stem *accommodate-*, and both trigger stem-level cycles: this is shown by the fact that the foot-head erected over the second syllable of *accómmodàte-* during the first cycle prevents the assignment of dactylic secondary stress to *accòmmodátion-* in the second cycle (compare again **àccommodátion* with *àbracadábra)*. The word-level phonology, in contrast, applies just once to the whole word.

While English only provides merely negative evidence, German affords a clear positive argument to show that the word-level phonology does not apply in cyclic domains smaller than the fully inflected grammatical word. In German, word-final consonants occupy the syllable coda at the word level, but resyllabify into the onset at the phrase level before enclitics beginning with a vowel: e.g. *spiel* [.ʃpiːl.] 'play' ~ *spiel es* [.ʃpiː.ləs.] 'play it' (Wiese 1996: 251; Hall 1999: 119). German coda devoicing must therefore be a word-level phonological process, since it overapplies to word-final consonants resyllabified before enclitics: e.g. *le*/g/ *es weg* [leː.kəs.vɛk] 'put it away'. In this light, consider how the word-level adjectival suffix *-ig* (Wiese 1996: 121) behaves with respect to devoicing. When *-ig* is the last overt suffix in the word, its final consonant is syllabified in the coda at the word level and so devoices in the normal way: e.g. *fett-ig* [fɛ.tɪç] 'fat-y'. In contrast, when *-ig* is followed by another word-level suffix beginning with a vowel, its final consonant undergoes resyllabification into the onset and, crucially, escapes devoicing: e.g. *fett-ig-es* [fɛ.tɪ.gəs] 'fat-y-N.NOM/ACC.SG'. This proves that the word-level phonology applies only once to *fett-ig-es*, even though this item contains two word-level suffixes (7,a). If *-ig* triggered a word-level cycle over the adjective stem, excluding the inflectional marker *-es*, devoicing would overapply (7,b).

(7)		a. correct domain structure	b. incorrect domain structure
		[$_{\text{WL}}$ *fett-ig-es*]	[$_{\text{WL}}$ [$_{\text{WL}}$ *fett-ig*] *es*]
$\mathcal{P}_{\text{WL}}$	1st cycle	fɛ.tɪ.gəs	fɛ.tɪç
	2nd cycle	–	*fɛ.tɪ.çəs

The nonrecursiveness of word-level domains also provides the key to a famous puzzle in Indonesian stress assignment.[9] In this language, suffixes are incorporated into the same prosodic word as the stem, whereas prefixes are prosodically adjoined (Cohn 1989: 200ff.):

(8)		
	($_{\omega}$ **bicára**)	'speak'
	($_{\omega'}$ məm ($_{\omega}$ **bicará-kan**))	'speak about'
	($_{\omega'}$ məm ($_{\omega}$ **bicàra-kán-ña**))	'speak about it'

Of interest here is the location of stress in polysyllabic sequences housed within the same prosodic word: in (8), these are highlighted in bold. When such sequences belong to a monomorphemic item, primary stress falls on the penult, and secondary prominence is assigned to every second syllable to its left. However, an odd-parity polysyllabic pretonic sequence, as in *xàtulistíwa*, will begin with a dactyl because the ω-initial syllable is required to bear stress and clash within ω is forbidden.

(9) bicára *bìcára 'speak'
màʃarákat 'society'
xàtulistíwa *xàtùlistíwa, *xatùlistíwa, *xatulistíwa 'equator'
èrodìnamíka 'aerodynamics'

As shown in (10), a stem formed with the suffix -(*n*)*isasi* '-ization' (< Dutch -*is-atie*) undergoes normal application of regular stress assignment, completely overwriting the metrical contour of its base. Note that the proparoxytonic contour of *amérika* is a lexical exception (Wallace 1976: 59).

(10) amérika 'America'
àmerikà-nisási *amèrika-nisási 'Americanization'

Crucially, -(*n*)*isasi* resembles the stem-level Latinate affixes of English: it entered Indonesian in words borrowed whole from Dutch, but it has since become an autonomous morpheme capable of attaching to new bases, including native stems. In such novel, productively derived forms, the allomorph -*isasi* is predictably selected when the base ends in a consonant, whereas -*nisasi* appears after vowels (De Vries 1984: 484–7; Mueller 2007: 1220–1).[10]

(11) a. kompor 'stove'
kompor-isasi 'the introduction of furnaces in brickyards'

b. pompa 'pump'
pompa-nisasi 'the introduction of pumping systems'

This confirms that the synchronically correct morphemic segmentation of *àmerikànisási* is as shown in (10).

In contrast with -(*n*)*isasi*, native suffixes like valency-changing -*kan* and 3SG.ACC -*ña* cause misapplication of secondary stress assignment.

(12) ($_{\omega}$ **bicára**) 'speak'
($_{\omega'}$ məm ($_{\omega}$ **bicará-kan**)) *($_{\omega'}$ məm ($_{\omega}$ **bìcará-kan**)) 'speak about'

In the derived applicative verb *məm-**bicará-kan***, primary stress falls on the penult, as normal. The preceding syllable, which heads a foot in the base ***bicára***, undergoes destressing to avoid clash: this too constitutes normal application. But the metrical profile of *məm-**bicará-kan*** is opaque because the first syllable of the stem (which is ω-initial) fails to receive secondary prominence, even though stressing it would not result in a clash. The reason is that the stem-initial syllable is unstressed in the base ***bicára***, and pretonic stress in words containing -*kan* and -*ña* is faithful to the cyclic base, save for the avoidance of clash. The same pattern can be observed in the derived causative verb *məŋ-**àmerikà-nisasí-kan*** (Cohn & McCarthy 1998: 66):

(13) ($_{\omega}$ **amérika**) 'America'
($_{\omega}$ **àmerikà-nisási**) 'Americanization'

($_{\omega'}$ məŋ ($_{\omega}$ **àmerikà-nisasí-kan**)) { *($_{\omega'}$ məŋ ($_{\omega}$ **àmerìka-nìsasí-kan**)) / *($_{\omega'}$ məŋ ($_{\omega}$ **amèrika-nisasí-kan**)) } 'Americanize'

In this case, the derived verb bears secondary stress exactly on the same syllables as its proximate base ***àmerikà-nisási***; the only difference is that the foot-head on the penultimate syllable of the proximate base disappears in the derived verb in order to avoid clash. As one would expect, moreover, *məŋ-**àmerikà-nisasí-kan*** is metrically faithful to its proximate base ***àmerikà-nisási***, rather than to its remote base ***amérika***.

We now come to the key puzzle: if a word contains a sequence of two native suffixes like -*kan* and -*ña*, it exhibits opaque stress too, but it is metrically faithful to its remote base, rather than to the proximate one.

(14) ($_{\omega}$ **bicára**) 'speak'
($_{\omega'}$ məm ($_{\omega}$ **bicará-kan**)) 'speak about'

($_{\omega'}$ məm ($_{\omega}$ **bicàra-kán-ña**)) { *($_{\omega'}$ məm ($_{\omega}$ **bìcara-kán-ña**)) / *($_{\omega'}$ məm ($_{\omega}$ **bicara-kán-ña**)) } 'speak about it'

In *məm-**bicàra-kán-ña***, primary stress falls on the penultimate syllable, as normal. The antepenult cannot bear secondary prominence, even though it corresponds to the tonic syllable of the proximate base *məm-**bicará-kan***, because this would create an illegal clash. Unexpectedly, however, pretonic stress in *məm-**bicàra-kán-ña*** falls on the second syllable of the stem. This metrical profile is unfaithful to the proximate base *məm-**bicará-kan***, where the second syllable of the stem is unstressed. It is also an opaque contour, in that the normal application of stress assignment would generate a stem-initial dactyl with prominence on the first syllable of the stem. The only explanation is that pretonic stress in *məm-**bicàra-kán-ña*** must reflect cyclic inheritance from the remote base ***bicára***. How can this be?

Stratal Phonology provides a simple solution to this classic puzzle: -(*n*)*isasi* is a stem-level suffix, but -*kan* and -*ña* are word-level. At the stem level, iterative stress assignment reapplies normally to each new cyclic domain, overwriting the metrical structure created in previous cycles. At the word level, in contrast, a single new foot is built noniteratively at the right edge, causing the penultimate syllable to receive primary stress; the antepenult is destressed if necessary to avoid clash, but otherwise the pretonic string remains undisturbed. Now, any form containing just one word-level suffix, like *məŋ-**àmerikà-nisasí***$_{\text{SL}}$***-kan***$_{\text{WL}}$ in (13), is predicted to be metrically faithful to its proximate base, as the latter defines an immediate cyclic subdomain: see (15,a). But forms containing two word-level suffixes in a row will behave differently: since the word-level phonology applies just once to the whole word, the first word-level suffix fails to trigger a cycle. Thus, an item like *məm-**bicàra-kán***$_{\text{WL}}$***-ña***$_{\text{WL}}$ in (14) ends up being metrically faithful to its remote base because the proximate base does not constitute a cyclic node; in this case, only the remote base defines an immediate cyclic subdomain, as shown in (15,b).

(15)	a. [$_{\text{WL}}$ məN [$_{\text{SL}}$ [$_{\text{SL}}$ amérika] (n)isasi] kan]	b. [$_{\text{WL}}$ məN [$_{\text{SL}}$ bicara] kan-ña]
SL: pre-σ́ overwriting	($_{\omega}$ amérika)	($_{\omega}$ bicára)
	($_{\omega}$ àmerikànisási)	
WL: pre-σ́ faithfulness	($_{\omega'}$ məŋ ($_{\omega}$ àmerikànisasíkan))	($_{\omega'}$ məm ($_{\omega}$ bicàrakánña))

Nothing else is needed to handle more complex cases like *məŋ-**àmerikà-nisàsi***$_{\text{SL}}$***-kán***$_{\text{WL}}$***-ña***$_{\text{WL}}$ 'Americanize it' (Cohn & McCarthy 1998: 72), with one stem-level suffix followed by two word-level suffixes:

(16)		[$_{\mathrm{WL}}$ məN [$_{\mathrm{SL}}$	[$_{\mathrm{SL}}$	amérika] (n)isasi] kan-ña]
	SL: pre-σ́ overwriting		($_{\omega}$	amérika)
			($_{\omega}$	àmerikànisási)
	WL: pre-σ́ faithfulness	($_{\omega'}$ məŋ	($_{\omega}$	àmerikànisàsikánña))

This item is faithful neither to its proximate base *məŋ-**àmerikà-nisasí-kan*** (where /sa/ is unstressed) nor to the most deeply embedded base ***amérika*** (which bears prominence on the second syllable). Rather, pretonic faithfulness targets the input to the word level: the noun stem ***àmerikà-nisási***. The result happens to exhibit the same metrical contour as a hypothetical transparently-stressed nine-syllable form, but a comparison with opaque word-level forms like (15) shows this outcome to be coincidental.

The stratal approach reveals a profound similarity between cyclic stress assignment in English and Indonesian. The English form *orígin-àt*$_{\mathrm{SL}}$-*ing*$_{\mathrm{WL}}$, with a stem-level suffix followed by a word-level one, is metrically faithful to its proximate base *orígin-àte*, and not to the remote base *órigin*. In contrast, *defénce-less*$_{\mathrm{WL}}$-*ness*$_{\mathrm{WL}}$, with its two word-level suffixes, inherits its metrical contour from its remote base *defénce*. Indonesian *məm-**bicàra-kán***$_{\mathrm{WL}}$***-ña***$_{\mathrm{WL}}$ reveals this shared pattern more clearly because, in this case, word-level noniterative refooting at the right edge forces the proximate base *məm-**bicará-kan*** to be less faithful to the remote base ***bicára***.

The behaviour of word-level suffixes, as illustrated by the evidence of German in (7) and of Indonesian in (15), raises difficulties for cyclic frameworks that posit richer domain structures than Stratal Phonology. In Cophonology Theory, for example, every nonterminal morphosyntactic node triggers a phonological cycle; a string of the form *base-affix*$_1$-*affix*$_2$ cannot define a single phonological cyclic domain unless its morphosyntactic structure is flat, i.e. [*base-affix*$_1$-*affix*$_2$] rather than [[*base-affix*$_1$]*affix*$_2$] (Orgun 1996: ch. 2). To accommodate data like (7), therefore, Cophonology Theory is forced to say that German suffixes like *-ig* are specified for a vacuous cophonology, i.e. $\mathcal{P}_{\text{-ig}}(i) = i$, and that coda devoicing is confined to a cophonology restricted to inflectionally complete grammatical words. For this solution to work more generally, however, it is not enough to ban all phonological unfaithfulness during vacuous cycles; phonologically driven allomorph selection must be suspended too, in order to permit outwards-sensitive phonologically conditioned allomorphy (e.g. Deal & Wolf 2017). Like the proliferation of zero morphs in mainstream DM (Bermúdez-Otero 2016: 392), the proliferation of vacuous cycles in Cophonology Theory does not constitute a direct refutation of the framework, but it could be taken as a signal that a generalization is being missed. Approaches to the interface based on Phase Theory face similar problems over the absence of phonological cyclic domains between the grammatical word and the utterance: Scheer (2011: §786) refers to this as the 'word spell-out mystery'.

5.2.3.3 The recursiveness of stem-level domains and Chung's Generalization

Rule-based Lexical Phonology used a large measure of brute force to handle the fact that stem-level domains are recursive whereas word-level domains are not (except in the rare cases mentioned in note 8). Affixes were arbitrarily labelled as stem-level or word-level, and it was stipulated that the former were cyclic and the latter were postcyclic or noncyclic (Booij & Rubach 1987; Halle & Vergnaud 1987).[11] This meant that each instance of stem-level affixation triggered a stem-level cycle, whereas only inflectionally complete grammatical words triggered word-level cycles. Whilst such provisions cover a remarkably large amount of empirical ground,[12] they are conceptually unsatisfying; it would be far better if the recursiveness of stem-level domains could, like the phonological inertness of roots, be deduced from independent postulates.

To this end, Bermúdez-Otero (2012: 19–20, 31–40; 2013b), developing intuitions adumbrated by Pesetsky (1979: §5.0) and Borowsky (1993: 219–20), suggests that the richer domain structure that characterizes stem-level constructs as compared with word-level forms emerges from facts about lexical decomposition and storage. His proposals make extensive use of Jackendoff's (1975) theory of lexical redundancy rules, adapted to the framework of OT. In this view, complex stem-level items like *accommodate* and *accommodation* are listed nonanalytically: i.e. as undecomposed wholes with all stem-level phonological properties redundantly specified (17,a). Of course, this in no way implies a denial of the psychological reality of stem-level morphology and phonology; it merely means that stem-level processes work as lexical redundancy rules. In contrast, complex word-level items like *accommodationlessness* may be unlisted or, if entered into the lexicon, are listed analytically as decomposed strings of stem-level pieces (17,b).[13]

(17) Stem-level nonanalytic listing

a. ACCOMMODATE ↔ accómmodàte

b. ACCOMMODATION ↔ accòmmodátion

Word-level analytic listing

c. ACCOMMODATIONLESSNESS ↔ [$_{\text{WL}}$ accòmmodátion -less -ness -∅]

In this system, stem-level domain recursion emerges from morphological blocking. Thus, when the noun *accommodation* was first created, the existence of a lexical entry for the verb *accommodate* (17,a) blocked derivation from the root *commod-*.[14] As this lexical entry already contained a foot-head on the second syllable, initial pretonic secondary stress was blocked too.

(18)

accómmodàte – ion	MAX-Head(Σ)	ALIGN(ω,L;Σº,L)
àccommodátion	*!	
accòmmodátion ☜		*

Since morphological blocking is affected by factors such as token frequency, this account correctly predicts that those factors will also have an effect on stem-level domain recursion (Collie 2007, 2008): for example, relative token frequency plays a key role in determining whether derivatives like *importation* and *transportation* cyclically preserve the foot-head on the second syllable of their bases (Bermúdez-Otero 2012: 34–9; 2013b: §21, §42–§45).

(19) tokens per million words in spoken section of COCA

				base		*derivative*
a.	imp[ɔ́]rt	~	ìmp[ɔ̀]rt-átion	5.15	>	0.62
b.	trànsp[ɔ́]rt	~	trànsp[ə]rt-átion	7.23	<	23.54

Like all versions of Stratal OT, this approach to stem-level domain recursion also explains Chung's Generalization (named after Chung 1983: 63). This states that, if a phonological process misapplies within an outer stem-level domain owing to the presence of an inner stem-level domain, then the output of that process must be lexically contrastive. For example, the faithfulness constraint that opaquely preserves second-syllable stress in the derivation of stem-level *accòmmodátion-* from stem-level *accómmodàte-* (18) also preserves the exceptional pretonic contour of monomorphemic *Epàminóndas* (20), which contrasts with that of monomorphemic *àbracadábra* (Kiparsky 2007: 20ff.; Bermúdez-Otero 2012: 31–3; Bermúdez-Otero 2013b: §32, §38–§41, and p. 22).

(20)

Epàminóndas	MAX-HEAD(Σ)	ALIGN(ω,L;Σº,L)
Èpaminóndas	*!	
Epàminóndas ☜		*

OT with OO-correspondence is unable to explain Chung's Generalization because it uses different constraints to describe phonemic contrast in simple items and cyclic effects in complex items (Kiparsky 2007: 22; Bermúdez-Otero 2012: 40).[15]

5.2.3.4 Cycles over word-level affixes

Our brief discussion of root inertness (§5.2.3.1) and of stem-level domain recursion (§5.2.3.3) has shown that, while the ideas of cyclic derivation and of stratification provide the stratal research tradition with an enduring and stable core, other aspects of the framework continue to be revised and improved while new empirical predictions are constantly being derived.

As a further illustration of ongoing work, let me return to the cyclic domain structure of word-level constructs (§5.2.3.2). In (7) and (15,b) we saw that word-level domains do not embed other word-level domains (except in the circumstances discussed in note 8). Beyond this, however, the stratal tradition affords two competing views of word-level affixes: one such view, represented in (21,a), holds that affixes never define cyclic domains (Kiparsky 1982a, 1982b); the other, exemplified in (21,b), holds that word-level affixes behave just like cyclic stems by defining domains for the stem-level phonology by themselves (Baker 2005: 17, developing ideas in Borowsky 1993).

(21) a. $[_{WL}\,[_{SL}$ memory] less-ness] b. $[_{WL}\,[_{SL}$ memory]$[_{SL}$ less]$[_{SL}$ ness]]

Baker's approach predicts that, whilst word-level affixes may *qua* functional morphs escape the prosodic minimality restrictions imposed by the stem-level phonology on lexical items, they will otherwise behave like miniature stems. Recent work supports this prediction with evidence from German and Dutch (Buckler 2009; Buckler & Bermúdez-Otero 2012). In German, for example, stem-level constraints require that an underlying voiceless dorsal fricative should be realized as [ç] in stem-initial position (22,a). Strikingly, the word-level diminutive suffix *-chen* behaves exactly like a stem, in that its initial dorsal fricative is mapped onto [ç] domain-initially at the stem level (22,b).

(22) a. **Ch**ina [çiːna] 'China'

b. $[_{WL}\,[_{SL}$ Kuh]$[_{SL}$ **ch**en]] $[_{SL}$ kuː-]$[_{SL}$ -çən] → $[_{WL}$ kuːçən] 'cow-DIM'
$[_{WL}\,[_{SL}$ Ku**ch**en]] $[_{SL}$ kuːxn̩] → $[_{WL}$ kuːxn̩] 'cake'

Baker's (2005: 17) hypothesis has further advantages: it naturally accounts for languages like Diyari (Poser 1989: 127–8) and Ngalakgan (Baker 2005: 4ff.) in which every productive suffix constitutes a separate stress-assignment domain, and it answers McCarthy's (2007: 133–4) question as to why the phonemic inventory of word-level affixes is never a superset of the phonemic inventory of stems. It has also proved helpful in analyses of nonconcatenative exponence in the framework of Generalized Nonlinear Affixation (§5.4.3): see e.g. Trommer (2011: 73ff.), Zimmermann (2016b: 273). In §5.4.2 below I shall put it to further use in a solution to bracketing paradoxes of the *ungrammaticality* type, as suggested by Baker (2005: 16–17).

In this light, the stratification generalizations in (6) should be amended as follows:

(23)
a. Roots do not define cyclic domains.
b. Some stems and some affixes define cyclic domains for the stem-level phonology ($\mathcal{P}_{SL}$).
c. Words define cyclic domains for the word-level phonology ($\mathcal{P}_{WL}$).
d. Utterances define cyclic domains for the phrase-level phonology ($\mathcal{P}_{PL}$).

5.3 In defence of Cyclic Containment

5.3.1 OO-correspondence

Like all cyclic approaches to the morphology–phonology interface, Stratal Phonology currently meets its most serious challenge in the theory of OO-correspondence (e.g. Kenstowicz 1996; Benua 1997). The key idea behind the latter may be summarily stated as follows; cf. (3) above.

(24) In cases of morphosyntactically induced phonological opacity, a linguistic expression copies its opaque phonological properties from the surface representation of a morphosyntactically related expression.

Under OO-correspondence, therefore, *accòmmodátionlessness* acquires its prefenestral primary stress and its nondactylic secondary stress from the surface representation of the singular form of the noun *accòmmodátion*, rather than from an intermediate representation assigned to the stem *accòmmodátion-* in the course of the derivation; cf. (2) above.[16]

Although in this specific instance both theories produce the same result, their wider predictions diverge dramatically. In particular, OO-correspondence holds that all opaque properties of the derived form must occur transparently in some surface base (Bermúdez-Otero 2011: 2029), and that surface bases need not correspond to morphosyntactic constituents of the derived form.

(25)

		Cyclic frameworks	**OO-correspondence**
a.	Need opaque properties surface transparently in the base?	NO	YES
b.	Need the base be contained within the derived form?	YES	NO

I shall briefly return to question (25,a) in §5.3.3 below. In §5.3.2 I look in depth at one piece of evidence bearing on question (25,b), which in effect asks whether Cyclic Containment

(3) is true or not; other data bearing on this issue are discussed in Bermúdez-Otero (forthcoming).

5.3.2 English dual-level affixes

Steriade (1999: §2–§3) challenges Cyclic Containment (3) with observations from English dual-level affixes like adjectival *-able*. Consider first the adjective *párodiable*, derived from the verb *párody*$_{V}$, itself formed by conversion of the noun *párody*$_{N}$: in this adjective, *-able* behaves as a stress-neutral suffix, creating an extremely long metrical lapse after the tonic syllable. Now suppose that we parse the adjective *remédiabl*e as derived through the addition of *-able* to the converted verb *rémedy*$_{V}$: in this analysis (to be revised presently), *-able* behaves as a stress-affecting suffix, causing primary stress to shift to the right. If so, what enables the stress shift in *remédiable*, whilst *párodiable* is forced to retain its relatively marked metrical profile? According to Steriade, it is the fact that the pre-existent adjective *remédial* provides a lexical precedent for the stress contour of *remédiable*, whereas there is no such precedent for **paródiable*: cf. $^{\varnothing}$*paródial.*[17] Steriade refers to this as an instance of 'lexical conservatism'.

The putative connection of *remédiable* with *rémedy*$_{V}$ and *remédial* is, however, questionable. In an alternative analysis, Raffelsiefen (2004: 135) regards the metrical profile of *remédiable* as licensed by that of the verb *remédiàte* (whence also *remèdiátion*). Even if the token frequency of *remédiàte* is lower than that of *remédial*, this proposal has the eminent virtue of subsuming the pair *remédiàte~remédiable* under a highly pervasive pattern linking verbs in *-ate* with adjectives in *-able* (Kiparsky 2005: 507). In turn, this pattern is clearly connected to the observation that, when added to stems, *-able* subcategorizes for verbs, not for adjectives. In relation to the ungrammaticality of **paródiable*, moreover, Kiparsky (2005: 507) points out that a lexical precedent for its stress pattern does exist in the adjective *paródic*. Yet, crucially, there is no verb $^{\varnothing}$*paródiàte* (and hence no noun $^{\varnothing}$*paròdiátion*). The corresponding token frequencies, measured in tokens per million words in the BNC, are as follows:

(26)	a.	rémedy$_{V}$	5.64	b.	párody$_{V}$	1.04
		remédial	3.39		paródic	0.43
		remédiàte	0.03		$^{\varnothing}$paródiàte	0
		remèdiátion	0.23		$^{\varnothing}$paròdiátion	0
		remédiable	**0.28**		**párodiable**	**0.01**

The metrical contrast between *remédiable* and *párodiable* thus constitutes a genuine instance of lexical conservatism, but the relevant lexical precedents seem to be *remédiàte* and $^{\varnothing}$*paródiàte*, not *remédial* and $^{\varnothing}$*paródial*.

Raffelsiefen's (2004) richly detailed study of verbs in *-ize* provides a more straightforward case of lexical conservatism in English dual-level affixation. The suffix *-ize* is highly productive in stress-neutral use: e.g. *cónsonant* → *cónsonant-ìze*. Raffelsiefen's data show, however, that *-ize* requires the immediately preceding syllable to be unstressed: e.g. *Clínton-ìze*, but **Búsh-ìze*. Crucially, this requirement can be met through stress shift, but typically only when there already exists another stress-shifted derivative from the same base:

(27)	a.	Japán	**Jàpan**-ése	**Jápan**-ìze	
		Vietnám	**Vìetnam**-ése	**Viétnam**-ìze	
	b.	Tibét	Tibét-an	*Tíbet-ìze	$^{\varnothing}$Tìbet-ése
		Brazíl	Brazíl-ian	*Brázil-ìze	$^{\varnothing}$Bràzil-ése

Evidence of this sort indicates that English dual-level affixes typically abide by the following generalization:

(28) Let *d* be an English derivative of the form *base+affix*.
Let *base* exist as a free form.
Let *affix* have a productive stress-neutral use.
If   *d* does not match the stress profile of *base*,
then   there exists another derivative *d′* of the form *base+affix′*
such that *d* matches the stress profile of *d′*.

In agreement with Raffelsiefen (2004: 135), I assume that the syntactic selectional restrictions of dual-level affixes constrain the availability of lexical precedents for stress shift under (28): this is shown by the fact that, as we saw in (26), *remédiable* is licensed by *remédiàte* rather than by *remédial*, and *paródic* fails to license **paródiable*. In what follows, however, I set this point aside and discuss Steriade's analysis in its original form: the goal of my argument is to show that, apart from the role of syntactic subcategorization, Steriade's account fails to capture an important phonological constraint on the stress-shifting uses of -*able*.

Steriade interprets generalization (28) as follows. If a lexical item like *remédial* occurs with sufficient frequency, its surface representation is stored in memory. The surface representations stored in memory are monitored by a set of optimality-theoretic Lex constraints. In the case of *remédiable* and *párodiable*, the relevant constraint is Lex-stress: for a candidate *c* containing a realization of a morpheme *m*, Lex-stress is violated if there is no stored surface realization of *m* containing the same sequence of stressed and unstressed syllables as the realization of *m* in *c*. Thus, *remédi-able* satisfies Lex-stress because it contains a surface realization of the root *remedy*$_{\surd}$- whose metrical properties match those of the realization of the root in *remédi-al*; in contrast, **paródiable* enjoys no such support.[18]

(29)

		Lex-stress	*Posttonic Lapse
parody-able *surface support*: párody	(pá.ro.)di.a.ble ☜		**
	pa.(ró.di.)a.ble	*!	*
	pà.ro.(dí.a.)ble	*!	
remedy-able *surface support*: rémedy, remédial	(ré.me.)di.a.ble		(*) *!
	re.(mé.di.)a.ble ☜		(*)
	rè.me.(dí.a.)ble	*!	

This analysis directly challenges Cyclic Containment (3): it claims that [ɹi.ˈmiː.di.ə.bl̩] copies its stress contour from the surface representation of a lexical item, *remedial*, that is not contained within *remediable*.

Kiparsky (2005) provides an alternative analysis of (28) in Stratal Phonology. According to his proposal, -*able* and -*ize* are dual-level suffixes, and their stratal affiliation depends on the morphosyntactic status of the base (Giegerich 1999). More specifically, a dual-level suffix like *-able* or *-ize* may occupy two different structural positions: it can attach 'high' to an inflectional stem, or it can attach 'low' to a root or derivational stem.[19] When the affix attaches high, it behaves as word-level and therefore stress-neutral; when it attaches low, it behaves as stem-level and therefore stress-affecting. Crucially, the default mode of attachment is high. Thus, *-able* in productive use normally combines with a verb's inflectional stem in word-level mode (e.g. *párody*$_{\mathrm{V}}$ → *párodi-able*); it can be added to a root in stem-level mode (e.g. *remedy*$_{\surd}$→ *remédi-able*, cf. *rémedy*$_{\mathrm{V}}$) only when the existence of a derived verb like *remédi-àte*$_{\mathrm{V}}$ alerts the learner to the availability of the root for stem-level suffixation. In this view, generalization (28) reflects facts about morphosyntax (the default attachment height for dual-level affixes) and about lexical acquisition (the learner's failure to extract roots in the absence of positive evidence), not about phonological computation.

The morphological premises of Kiparsky's analysis are strongly supported by crosslinguistic evidence. First, low attachment is known to be strongly correlated with reduced productivity, noncompositional semantics, and arbitrary allomorphy (e.g. Arad 2003). As a special case, De Belder, Faust & Lampitelli (2014) show that low diminutives are far less productive than high diminutives. Secondly, there are languages in which all synchronically productive derivation is based on inflectional stems: examples include Spanish (Bermúdez-Otero 2013a) and Portuguese (Matzenauer & Bisol 2016), where inflectional stems can be formally recognized by the presence of a final theme vowel.[20] In the case of Spanish, Bermúdez-Otero (2013a) adduces internal evidence from local domains for allomorph selection and external evidence from response latencies in lexical recognition to show that Spanish speakers store full inflectional stems in the lexicon. Thus, Spanish speakers' preferences in respect of derivational bases and of lexical decomposition closely match the learner behaviour assumed in Kiparsky's explanation of lexical conservatism in English dual-level affixation.

Kiparsky's analysis is accordingly able to derive observation (28) from independently motivated postulates. Because it assumes a stratal architecture, however, it also makes further empirical predictions. Notably, Kiparsky holds that the stress-shifting uses of dual-level affixes with independently surfacing bases, as in (27,a), reflect the application of the stem-level phonology. Under the stratification generalizations set forth in (23), however, the stem-level phonology applies in all cycles triggered by stems, including those stems created by derivation from a bound root. The stratal account therefore yields the following prediction:

(30) The stress-shifting uses of an English dual-level suffix follow the same pattern of primary stress assignment as items formed by adding the same suffix to a bound root.

This is just a special case of (31).

(31) A form created by adding a stem-level affix to a free stem obeys the same well-formedness conditions as a form created by adding the same suffix to a bound root.

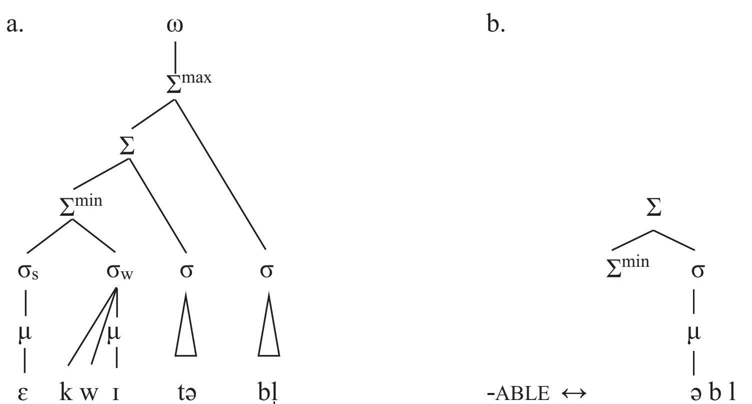
Consider the case of *-able*. To test the validity of (30), we need to examine stress assignment in adjectives formed by adding *-able* to bound roots that have no other derivatives; by ensuring that the root does not occur elsewhere, we guarantee that lexical conservatism is not at work. If the stratal analysis is correct, then the phonological generalizations governing stress assignment in these isolated deradical adjectives should correctly describe the behaviour of primary stress in the stress-affecting uses of -*able* with independently surfacing bases.

The data in (32) reveal that, when attached to bound roots, the suffix -*able* behaves like a weak retractor: it places primary stress on the immediately preceding syllable, i.e. on the antepenultimate, if heavy (32,a); otherwise, stress goes on the preantepenultimate (32,b).[21] For a survey of stress retraction modes in English, see Kager (1989: 37–63) and references therein.

(32)	a.	*heavy antepenult*	b.	*light antepenult*
		amḗnable		indómĭtable
		coméstible		indúbĭtable
		deléctable		inéxŏrable
		inelúctable		irréfrăgable

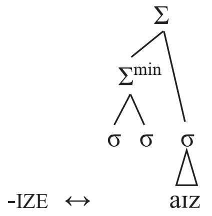
Weak retraction is not the default metrical pattern for disyllabic stem-level suffixes, as indicated by the prefenestral stresses in (32,b): cf. e.g. -*ity*. *Pace* Raffelsiefen (2004) and Zamma (2012), however, it is by no means necessary to set up a separate stem-level cophonology for weak retractors. In fact, doing so would lead to grievous loss of generalization, as repeatedly pointed out in the literature (e.g. Hayes 1982: 243). Rather, since it is independently known that the stem-level phonology of English achieves exhaustive footing by means of adjunction (Bermúdez-Otero 2012: footnote 19 and references therein), a weak-retraction structure like (33,a) can be obtained simply by specifying the first syllable of -*able* as sister to a minimal foot projection ($\Sigma°$ or Σ^{min}; see Ito & Mester 2009: 170) in the underlying representation of the suffix (33,b). Faithfulness to this specification will be enforced by the high-ranking constraint IDENT-$\sigma\frown\Sigma°$.

(33)



Once this mechanism is in place, the same stem-level constraints that build right-aligned bimoric trochees with final syllable adjunction in monomorphemic items like *a*(*gén*)*da* and *A*(*méri*)*ca* will generate weak retraction in *ine*(*lúc*)*table* and *in*(*dómi*)*table.*[22] Strong retractors like -*ize*, which cause stress to skip the immediately preceding syllable regardless of weight (e.g. *ré**cog**nìze*), yield to the same analysis, except that the suffix is underlyingly specified as adjoined to a syllabic trochee:

(34)



We can now verify prediction (30). As expected, *remédiable* complies with the pattern of weak retraction found in isolated deradical adjectives in *-able* (32): its canonical pronunciation is [ɹi.ˈmiː.di.ə.bl̩], with preantepenultimate stress before a light antepenult.[23] Similarly, the metrical contour of verbs containing *-ize* in stress-shifting use, like *Jápanìze* and *Viétnamìze* in (27,a), is consistent with the pattern of strong retraction exhibited by deradical items like *récognìze*.

Under the analysis in (29), these facts come as a surprise: Lex-stress predicts that derived adjectives in *-able* will adopt the least marked metrical configuration for which a lexical precedent is available; there is no provision to ensure that this configuration will always be one of weak retraction, nor that it will always match the behaviour of isolated deradical formations like those in (32). This problem goes beyond mere loss of generalization: Lex-stress incorrectly predicts that *-able* will shift primary stress onto a light antepenultimate syllable, in violation of the weak-retraction pattern, whenever some established derivative of the same root provides a precedent for antepenultimate stress and none supports preantepenultimate stress.

Consider, for example, the verb *period*$_{V}$ [ˈpɪə.ɹi.əd], derived by conversion from the homophonous noun. The verb provides the base for the low-frequency novel adjective *period-able*.[24] Since the stress-neutral pronunciation *périodable* [ˈpɪə.ɹi.ə.də.bl̩] contains a long final lapse, Lex-stress predicts that a candidate with rightward stress shift will be preferred if supported by some established form containing the root *period*$_{\surd}$. There happen to be two such established forms: *pèriódic* and *pèriódical*. In consequence, the grammar in (29) generates **pèriódable*: this candidate has a precedent in established derivatives of *period*$_{\surd}$ and avoids a long final lapse. Yet native speakers judge **pèriódable* to be completely ungrammatical. On top of the fact that, as adjectives, *periodic* and *periodical* belong to the wrong category in terms of syntactic selection, there is a clear problem on the phonological side too: a stress-shifting use of *-able* must conform with the same pattern of weak stress retraction as found in *-able* derivatives from bound roots (32), but **pèriódable* fails to do so.[25] Kiparsky (2005) runs through several variations of this argument.

Crucially, this empirical problem remains even if the Lex-stress analysis is supplemented with an explicit account of weak retraction in isolated deradical forms like (32). Suppose that, as proposed above, *-able* is underlyingly specified with a metrical marking requiring it to be footed by adjunction (33,b), and that compliance with this specification is monitored by the faithfulness constraint Ident-σ⁀Σ°. This constraint will need to be ranked above *PosttonicLapse in order for the metrical specifications of *-able* to induce prefenestral stress in isolated deradical items: in tableau (35), this is illustrated with the derivation of preantepenultimate stress in *indómitable*. In a monostratal analysis, however, Lex-stress must dominate Ident-σ⁀Σ° so as to account for departures from weak retraction in stress-neutral uses of *-able*: in tableau (35), this is the case of *párodiable*. These ranking arguments lead to the hierarchy Lex-stress ≫ Ident-σ⁀Σ° ≫ *PosttonicLapse, which correctly preserves the original outcome of the analysis for *remédiable*: support from *remédial* allows *remédiable* to satisfy the metrical specifications of *-able*. We have now expanded tableau (29) to handle *indómitable*, while preserving its results for *párodiable* and *remédiable*. As noted in the preceding paragraph, however, the system still selects the wrong candidate for *period-able*: **períodable* loses for lack of precedents in other derivatives of *period*, and **pèriódable* wins over *périodable* on low-ranking *PosttonicLapse.[26]

(35)

		Lex-stress	Ident-σ⁀Σ°	*Posttonic Lapse
indomit- $_{\Sigma^\circ}$)able	in.(dó.mi.)ta.ble ☞	(*)		*
surface support: ∅	ìn.do.(mí.ta.)ble	(*)	*!	
parody- $_{\Sigma^\circ}$)able	(pá.ro.)di.a.ble ☞		*	**
surface support: párody	pa.(ró.di.)a.ble	*!		*
	pà.ro.(dí.a.)ble	*!	*	
remedy- $_{\Sigma^\circ}$)able	(ré.me.)di.a.ble		*!	**
surface support: rémedy, remédial	re.(mé.di.)a.ble ☞			*
	rè.me.(dí.a.)ble	*!	*	
period- $_{\Sigma^\circ}$)able	(pé.ri.)o.da.ble ☹		(*)	**!
surface support: périod, periódic	pe.(rí.o.)da.ble	*!		*
	pè.ri.(ó.da.)ble 💣		(*)	

Stratal OT does not incur this problem. At the stem level, Ident-σ⁀Σ° is ranked above all constraints penalizing stress retraction; at the word level, it is inactive because it is crucially dominated by Ident-stress. At the stem level, therefore, *-able* can only behave as a retracting suffix, and at the word level it can only be stress-neutral. **Pèriódable*, which follows neither pattern, can win at neither level, regardless of any hypothetical support from *pèriódic*.

In sum, Kiparsky's analysis of English dual-level affixes derives generalization (28) from independently motivated premises, avoids the incorrect predictions of the Lex-stress constraint (e.g. the failure of retraction in **pèriódable*), and explains further facts (e.g. the parallelism between *remédĭable* and deradical *indómĭtable*). I conclude that lexical conservatism in stress-affecting uses of English dual-level affixes is a real and important empirical phenomenon, but one that is perfectly compatible with Cyclic Containment (3). Stratal Phonology explains its fine-grained detail better than OO-correspondence. Bermúdez-Otero (forthcoming) reaches the same conclusion about other instances of lexical conservatism alleged to challenge Cyclic Containment (cf. Steriade 1999, 2008).

5.3.3 Theory comparison

As the preceding case study shows, the question whether phonological derivations proceed cyclically is strictly empirical: the answer depends on the validity of generalizations such as (3) and (4). While the debate will no doubt continue in years to come, Stratal Phonology emerges from this challenge as a progressive research programme (Lakatos 1970) in that it responds to tough empirical tests not by weakening its empirical content but by producing results like (30).

The evolution of the theory of OO-correspondence since its birth more than twenty years ago (Benua 1995) looks rather different. From the outset, the practitioners of OO-correspondence have postulated a widening range of transderivational relationships between surface forms. As a result, we can now choose between asymmetrical and symmetrical correspondence, between local and nonlocal relationships, and between reference to free bases only and reference to all paradigmatically related expressions: compare, for example, the proposals of Benua (1997) and Kager (1999) with those of Burzio (1996) and Kenstowicz (1996). By itself, this growth in the number of transderivational correspondence types need not be a worrying sign; after all, one often sees a similar expansion and diversification of applications whenever a new grammatical mechanism of some generality is discovered. The problem lies, rather, in the fact that this trend has not been accompanied by the formulation of criteria defining the situations in which each type of transderivational relationship holds. Insightfully, Kenstowicz (1996: 390–1) identified this as an urgent task for the theory of OO-correspondence. Yet, to date, no such criterion has stood up to scrutiny. For example, McCarthy (2005: 172) proposed that OO-correspondence is asymmetrical and base-prioritizing in derivation, but symmetrical in inflection; but Hall & Scott (2007) and Albright (2008), among others, soon identified counterexamples. The outcome is that, as the generative capacity of the theory has grown in line with the range of its applications, its predictive power has fallen: if one has to predict in order to explain (Hempel & Oppenheim 1948), OO-correspondence describes more and more, but explains less and less.

As the empirical content of the theory of OO-correspondence dwindles, it becomes proportionately more difficult to find direct empirical counterexamples. In (25,a), however, we saw that all forms of transderivational correspondence converge on one prediction: OO-correspondence cannot explain phonological opacity in a linguistic expression unless its opaque phonological properties surface transparently in some morphosyntactically related form. Accordingly, OO-correspondence is directly falsified by all cases of morphosyntactically induced misapplication where no appropriately related surface form is transparent. Recent research has identified no fewer than seven instances of this phenomenon:

(36)
a. Schwa epenthesis in Itelmen intransitive verbs (Bobaljik 2008).
b. Voicing of word-final prevocalic /s/ in Ecuadorian Spanish (Bermúdez-Otero 2011: §6; Strycharczuk, Veer, Bruil & Linke 2014).
c. Lenition of linking and intrusive [ɹ] in nonrhotic English dialects (Bermúdez-Otero 2011: §7).
d. Stress in Albanian deponent verbs and *plurale tantum* nouns (Trommer 2013).
e. Debuccalization of word-final prevocalic /s/ in Northern Chilean dialects of Spanish (Broś 2015: ch. 4).
f. Failure of gliding of stem-final prevocalic /i/ in Bothoa Breton verbs (Iosad 2017: ch. 10).
g. Failure of vowel reduction in Catalan compounds (Mascaró 2016).

5.4 Morphological implications of Stratal Phonology

In §5.3.2 we saw Stratal Phonology meeting a new observational challenge without loss of empirical content. We must now ask, however, whether the theory achieves its success by placing unreasonable demands on morphology and syntax. Linguistic theory runs a constant risk of delivering illusory advances in the study of one part of grammar by smuggling the analytic costs across the interface with another component.

Recent developments in morphology illustrate this danger. Research within DM has uncovered robust and profound empirical generalizations about locality restrictions on suppletive allomorphy (Bobaljik 2012; Smith et al. 2016). However, mainstream versions of DM with vocabulary insertion into single terminals (e.g. Embick 2010) have produced theories of allomorphic locality that are demonstrably too restrictive (Merchant 2015; Bermúdez-Otero 2016). Counterexamples are typically avoided by shifting the burden onto phonology in ways that deprive phonological theory of its empirical content (Bermúdez-Otero 2013a: 87–91, 2016: 404–19).

In this section I argue that Stratal Phonology, in contrast, permits a graceful integration with morphology: it can derive the relative ordering of phonological strata without recourse to the Affix Ordering Generalization (§5.4.1), it can handle bracketing paradoxes without recourse to rebracketing operations (§5.4.2), and it favours restrictive approaches to apparently nonconcatenative exponence (§5.4.3).

5.4.1 Affix order

Stratal Phonology is often claimed to make untenable assumptions about morphology (e.g. Inkelas 2012: 157; Shwayder 2015: 42–4). The argument runs as follows: in Stratal Phonology, the serial order of phonological levels crucially depends on the Affix Ordering Generalization, which holds that all stem-level affixes must occur inside all word-level affixes (Selkirk 1982: 91, after Siegel 1974 and Allen 1978); but the Affix Ordering Generalization is false (Aronoff 1976: 85, Aronoff & Sridhar 1983, Fabb 1988, etc.), and so Stratal Phonology must be wrong.

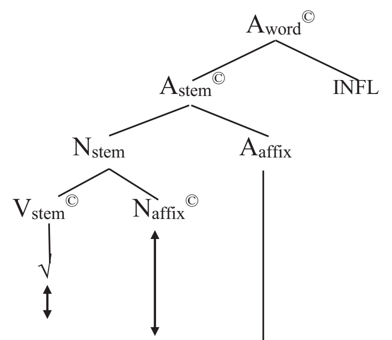
This argument, if correct, would indeed deprive Stratal Phonology of much of its appeal. Notably, one of the major advantages of the theory is its ability to derive phonological opacity effects from the size of the cyclic domains of the phonological processes involved (Kiparsky 2000, 2015b). This result supports a promising approach to the difficult problem of explaining the acquisition of opaque phonological derivations (Bermúdez-Otero 2003; Tesar 2014: 170–1, 399). In general, however, stratal accounts of opaque phonology would be thrown into disarray if violations of the Affix Ordering Generalization could disrupt the serial sequence of phonological levels by causing the word-level phonology to apply before the stem-level phonology.

Worryingly, the minor premise of the argument appears to be true: the Affix Ordering Generalization does appear to be untenable. Kiparsky (1983) rejected some of the putative counterevidence from English by arguing that, in cases like *cànnibalístic*, the nonexistence of $^{\varnothing}$*cánnibalist* indicates that the string -*istic* should be analyzed as a fused stem-level suffix, rather than as word-level -*ist* followed by stem-level -*ic*. Kiparsky (1983) also sought to deflect counterexamples like *devèlop-mént*$_{\mathrm{WL}}$-*al*$_{\mathrm{SL}}$ by invoking the dual-level status of -*ment*: he argued that, in fact, *devèlopméntal* has the same structure as *òrna-mént*$_{\mathrm{SL}}$-*al*$_{\mathrm{SL}}$, where -*ment* behaves as a stem-level affix because it attaches to a bound root (see §5.3.2 above). There is a valuable insight behind this suggestion: Hay (2003) has demonstrated experimentally that the acceptability of a novel item containing *-al* attached to a base ending in *-ment* decreases in direct proportion to the decomposability of the base. Nonetheless, Hay's experiments also show that, when coerced to add *-al* to an unequivocally word-level form ending in *-ment*, native speakers of English have no difficulty computing its phonological form: e.g. *impòverish-mént-al* [ɪmˌpɒvəɹɪʃˈmɛntl̩]. More decisively, the Affix Ordering Generalization appears not to hold in certain languages other than English (Inkelas 2012).

Crucially, however, the major premise of the argument from affix ordering is false: Stratal Phonology need not assume the Affix Ordering Generalization in order to prevent the word-level phonology from applying before the stem-level phonology; the stratification generalizations in (23) suffice to do the job. For the purposes of demonstration, let us assume, *pace* Kiparsky (1983), that *-ment* behaves as a word-level affix in *devèlop-mént*$_{WL}$*-al*$_{SL}$. The vital point is that, even if *developmental* has this structure, the suffix *-al* undeniably attaches to a noun stem (5,b), and not to a noun wordform (5,c), since there is no number inflection inside *-al*: plural *development-s* is fine, but word-based **development-s-al* is not (cf. note 8). It follows that, insofar as *development-al* is a stem-based derivative, and only fully inflected grammatical words trigger cycles of the word-level phonology (23,c), there is no word-level cycle before the addition of *-al* in the phonological derivation of *developmental*. Rather, given the syntactic structure in (37,a) and its morphological realization in (37,b), the theory of stratification summarized in (23) yields the composite phonological function in (37,c).[27] This results in a derivation in which all stem-level cycles precede all word-level cycles.

(37)

a.



b. [[[develop] [ment$_{WL}$] al$_{SL}$]]

c. $\mathcal{P}_{WL}$($\mathcal{P}_{SL}$($\mathcal{P}_{SL}$(develop), $\mathcal{P}_{SL}$(ment), -al))

5.4.2 Bracketing paradoxes

The term *bracketing paradox* is considerably vague and denotes a disparate collection of phenomena (Spencer 1988: 680–1; Stump 1991: note 38).[28] It is usually applied to problems of morphological analysis arising from a clash between two or more criteria for determining the constituent structure of a linguistic expression, but the conflicting criteria may be of various sorts: e.g. semantic scope, subcategorization requirements, syntactic distribution, etc. In this section I shall consider just one type of paradox exemplified by the English word *ungrammaticality*, which is widely believed to raise particular difficulties for Stratal Phonology; for discussion of *transformational grammarian*, see Bermúdez-Otero (2016: 422–3).

In general, bracketing paradoxes are challenging because, as we saw in (2), Stratal Phonology derives the order of cycles in the phonological derivation from part-whole relationships in morphosyntactic constituent structure. When morphosyntactic constituency and phonological domains seem not to match, a problem arises. Such challenges can always be overcome by brute force, i.e. by resorting to *ad hoc* rebracketing operations (Sproat 1985: 79ff., 468–9); but expedients of this nature cause the theory to haemorrhage empirical content.

Let us therefore consider the case of *ungrammaticality*. The Siegel–Allen theory of affix ordering (which I rejected in §5.4.1) suggests the morphosyntactic bracketing in (38,a). In contrast, semantic scope favours (38,b), since *ungrammaticality* means 'property of being ungrammatical'.

(38) a. [$_{\text{N}}$ un [$_{\text{N}}$ [$_{\text{A}}$ grammatical] ity]]
b. [$_{\text{N}}$ [$_{\text{A}}$ un [$_{\text{A}}$ grammatical]] ity]

The correct structure is (38,b), as shown by considerations of subcategorization: the prefix *un-* does not attach productively to nouns, at least not with the relevant meaning. Sproat (1985: 25–33) provides a detailed critique of Allen's (1978) and Fabb's (1984) claims that denominal *un-* prefixation is productive. Some apparent exceptions, like the well-established word *unbelief*, could plausibly be analyzed as diachronic back-formations from the corresponding adjectives (i.e. *unbelieving*). Other items formed by denominal *un-* prefixation appear to vary in their acceptability: for example, both *unproblem* and *unidiom* are attested on the World Wide Web but fail to occur in controlled corpora such as the BNC. Recently, Horn (2005) has argued that *un-* does in fact productively attach to nouns, but he crucially shows that the lexical semantics and pragmatics of the resulting words can be very different from the usual 'opposite' sense of *un-* in deadjectival items (e.g. *unhappy*) or the usual 'reverse' sense of *un-* in deverbal items (e.g. *unfasten*). In *ungrammaticality*, we do have the normal 'opposite' sense associated with *un-* in deadjectival derivation.

However, if the morphosyntactic constituent structure of *ungrammaticality* is that shown in (38,b), a serious phonological problem arises. The prefix *un-* is in the scope of the stem-level phonological cycle triggered by the suffix *-ity*. If so, what prevents the final consonant of the prefix from undergoing nasal place assimilation in this cycle, yielding *[ˌʌŋɡɹəˌmætɪˈkælətɪ]? Observe that, at the stem level, nasal+plosive clusters are subject to obligatory assimilation within feet (Kiparsky 1979: 439–40), as shown by their behaviour both in tautomorphemic environments (e.g. *co**n**g**a* [ˈkɒŋɡə]) and across the boundary between an affix and a bound root (e.g. *co**n**-**g**reg-ate* [ˈkɒŋɡɹɪˌɡeɪt]).[29]

Baker's (2005: 16–17) theory of word-level affixation, illustrated in (21,b) above, offers a simple solution, drawing on insights from Aronoff & Sridhar (1983) and Booij & Lieber (1993). Since *un-* is a word-level prefix, it defines a stem-level domain all by itself. But we know independently that every English word-level prefix occupies a prosodic word (ω) by itself too (e.g. Booij & Rubach 1984: §4.1). We can therefore infer that this prosodic word is erected over the prefix during its stem-level cycle. If this reasoning is correct, the phonological derivation of *ungrammaticality* runs as follows:

(39) *ungrammaticality*

	Domains		[$_{\text{WL}}$ [$_{\text{SL}}$ [$_{\text{SL}}$ ʌn][$_{\text{SL}}$ ɡɹæmætɪkæl] ɪtɪ]]
	$\mathcal{P}_{\text{SL}}$	1st cycle	($_{\omega}$ ˈʌn), ($_{\omega}$ ɡɹæˈmætɪkæl)
		2nd cycle	($_{\omega'}$ ($_{\omega}$ ˌʌn)($_{\omega}$ ɡɹæˌmætɪˈkælɪtɪ))
	$\mathcal{P}_{\text{WL}}$		($_{\omega'}$ ($_{\omega}$ ˌʌn)($_{\omega}$ ɡɹəˌmætɪˈkælətɪ))

Crucially, the prosodic word erected over *un-* in its affix cycle protects the final consonant of the prefix from place assimilation in the cycle triggered by *-ity*, as /n/ and /ɡ/ occupy different feet.[30]

Interestingly, Stratal Phonology offers an independent check on the validity of this solution. The prosodic word erected over *un-* in its affix cycle can survive in the cycle triggered by *-ity* only if preserved by a high-ranking faithfulness constraint in the stem-level phonology of English. Let us suppose that this constraint is IDENT-Head(ω), which requires that, if a foot heads a prosodic word in the input, it should also do so in the output. By the logic of Chung's Generalization, outlined in §5.2.3.3, this predicts that ω-headship must be lexically contrastive in English monomorphemic words. This proves correct: compare the regular prominence relationships in *bàrracúda* and *rèferéndum*, where the second foot heads the prosodic word, with the exceptional metrical pattern of *Ládefòged* and *périwìnkle* (Liberman & Prince 1977: 270, 308).

5.4.3 The division of labour between morphology and phonology

I conclude with a very small note on a very large topic. In the preceding sections we have seen that Stratal Phonology does not burden morphological theory with untenable assumptions: it can do its appointed job of deriving cyclic domains for phonological processes from morphosyntactic structure without having recourse either to level ordering or to rebracketing operations. More fundamentally, however, Stratal Phonology does presuppose the existence of a determinate boundary between morphology and phonology (Bermúdez-Otero 2012: 72; cf. Inkelas 2012: §4). This is because, in practical terms, to adopt the hypothesis of stratification is to assert that each pattern of exponence observed on the surface can be decomposed into a morphosyntactic and a phonological component (either or both of which may be vacuous). In turn, the morphosyntactic operations identified by such decomposition will fall into a small number of classes (known as 'levels' or 'strata') according to their phonological effects. If the theory of stratification summarized in (23) is correct, there are precisely three such classes.

Practising Stratal Phonology therefore requires robust and principled criteria for demarcating morphology from phonology. The best conceptual motivation for such a demarcation is to be sought in the principle of representational modularity (Jackendoff 1997: §2.6): morphology performs computations over morphs, whereas phonology performs computations over melodic and prosodic units. In this approach, morphology is bound by the Morph Integrity Hypothesis (Bermúdez-Otero 2012: 50): morphological operations may only select and insert exponents without altering their phonological content; 'process morphology' is banned. This conclusion is admittedly controversial: it is notably rejected in much work that otherwise shares key assumptions with Stratal Phonology (e.g. Anderson 1992; Inkelas 2012). Its great advantages, however, lie in its empirical content and heuristic power.

In line with the Morph Integrity Hypothesis, there is a long tradition of research that seeks to reduce apparently nonconcatenative exponence to the insertion of pieces of nonlinear phonological representation whose existence is independently motivated: e.g. floating features or feature-geometric treelets in the case of mutation, fully or partially bare prosodic nodes or prosodic treelets in the case of reduplication and subtraction. The general programme, pioneered by Lieber (1992: ch. 5) and Stonham (1994), is labelled Generalized Nonlinear Affixation in Bermúdez-Otero (2012: 53). Recent contributions to this line of research include Artés (2016), Bye & Svenonius (2012), Gribanova (2015), Iosad (2014), Köhnlein (2016: §5.1), Oostendorp (2012), Roseano (2015), Saba Kirchner (2010, 2013), Spahr (2016), Trommer (2011, 2014, 2015), Trommer & Zimmermann (2014), Zdziebko (2015), and Zimmermann (2013a, 2013b, 2016a, 2016b, 2017), among others.

5.5 Summary

The roots of Stratal Phonology are ancient (§5.2.1), and its conceptual core, consisting of the principles of cyclicity (§5.2.2) and stratification (§5.2.3), displays remarkable stability; but the theory also continues to grow and develop around this core. A promising line of research seeks to increase the explanatory depth of Stratal Phonology by deducing previously established generalizations, like the noncyclic status of roots and the recursiveness of stem-level domains, from independent facts (§5.2.3.1, §5.2.3.3). We have also seen that the theory responds to the challenge of new observations like lexical conservatism (§5.3.2) not by reducing its empirical content, but by producing predictions like (30). Crucially, Stratal Phonology achieves these results without imposing unreasonable theoretical costs on morphology: it can thrive without the Affix Ordering Generalization (§5.4.1) and without rebracketing operations (§5.4.2), and it has an intrinsic affinity for restrictive approaches to apparently nonconcatenative exponence like Generalized Nonlinear Affixation (§5.4.3). It is for these reasons that the stratal tradition continues to exert a claim on our attention.

5.6 Further reading suggestions

The classic exposition of rule-based Lexical Phonology is Kiparsky (1982b). Kaisse & Shaw (1985) provide a very accessible introduction. The relationship of the theory with Prosodic Phonology was elucidated by Booij & Rubach (1984), and Bermúdez-Otero & Luís (2009) provide up-to-date discussion of this question. Booij & Rubach (1987) codified the canonical three-level version of the framework. By the early 1990s, however, rule-based Lexical Phonology was in crisis, largely as a result of problems raised by the principles that sought to regulate the application of rewrite rules at the stem level: this situation is documented in Hargus & Kaisse (1993).

The most influential early presentation of Stratal OT, focused on the problem of opacity, is Kiparsky (2000); a more recent review of the *état de la question* may be found in Kiparsky (2015b). Bermúdez-Otero (2010) provides a comprehensive and continuously updated survey of Stratal OT, with reading suggestions and links, but referring only to my own work. Defences of Stratal Phonology against OO-correspondence include Bermúdez-Otero (2011, forthcoming), Trommer (2013), and Kiparsky (forthcoming); see also (36) above. Bermúdez-Otero (2012) presents a stratal perspective on the division of labour between phonology, morphology, and the lexicon.

As noted in §5.1, Stratal Phonology has consequences for many issues: on nonconcatenative morphology in general, see Bermúdez-Otero (2012: §2.4.2); on mutation morphology in particular, see Trommer (2011); on reduplication in particular, see Kiparsky (2010); on allomorphic locality, see Bermúdez-Otero (2013a: §3, 2016); on the interface with phonetics, see Bermúdez-Otero (2015: §22.2.2, §22.2.4) and Ramsammy (this volume); on phonological acquisition, see Bermúdez-Otero (2003), Boersma & van Leussen (2017), and Nazarov & Pater (2017); on phonological variation, see Turton (2016); and on historical change and the life cycle of sound patterns, see Bermúdez-Otero (2015), Kiparsky (2015a), and Ramsammy (2015, this volume).

Analyses of extensive fragments of the phonology of a single language or of closely related languages provide a good way to appreciate the heuristic value and explanatory power of stratal models. Classics in rule-based Lexical Phonology include Rubach (1984, 1993). In Stratal OT, see Kiparsky (2003a, 2003b) and Iosad (2017).

Notes

1 This term is due to Steriade (2013, forthcoming).
2 See Archangeli & Pulleyblank (this volume) for related ideas.
3 These principles build on a more general assumption of modularity: for discussion, see Bermúdez-Otero (2012: §2.4, 2015: §22.2) and §5.4.3 below.
4 In cases of phonological variation an input is associated with more than one output. If so, phonology specifies relations rather than functions (Smolensky 2006: 535–6; see also Kaye 1995: 330, note 18), and the cycle is more properly described as involving the composition of relations.
5 Stratal Phonology is compatible with a very broad range of approaches to word syntax. The specifics will not be crucial here. For related remarks, see Bermúdez-Otero (2012: note 38, 2013a: note 36).
The word *accommodationlessness* is naturally attested in a range of broadly compositional senses reflecting the structure in (2,a): for example, it occurs in reference to the medical condition in which the eye lacks the ability to perform the task of accommodating to the distance of visual objects.
The terms *root*, *stem*, and *word*, as used in (2,a), are defined in (5) below.
6 It does not matter whether the underlying representation of the adjective *long* has been restructured and no longer contains a final /ɡ/. The crucial point is that stem-final [ŋɡ] cannot occur before suffixes like -*ish* and that this phonotactic restriction on word-level derivatives is opaque (Bermúdez-Otero 2011: 2020, footnote 2).
7 These correspondences are only approximate, however. In Lexical Phonology, cyclic rules were assumed to abide by Strict Cyclicity and Structure Preservation (see §5.2.1 above), and these hypotheses were often maintained in the face of disconfirming evidence by assigning a rule to the postcyclic stratum even though it applied in domains smaller than the word (Bermúdez-Otero 2013b: §23–§29).
8 Word recursion only arises in the intended sense when a fully inflected grammatical word acts as the base for the derivation of a new stem, which is then itself inflected. This is not impossible: see Rainer (1996) and Bermúdez-Otero (2013a: 26) for examples.
9 My Indonesian data come from Cohn (1989) and Cohn & McCarthy (1998). I avoid examples containing schwa, which Cohn (1989: 174) describes as metrically invisible. Goedemans & van Zanten (2007) have recently argued that, in fact, Indonesian has no word stress at all. The case may be similar to that of Spanish secondary stress: there is no direct acoustic manifestation of its existence, but Hualde & Nadeu (2014) find subtle indirect evidence that supports the footing pattern implicit in traditional reports.
10 There is some variation: the allomorph -*nisasi* can occur after /r/, and bases ending in /i/ sometimes take -*sasi*.
11 Affixes can also display dual affiliation, in which case their phonological behaviour correlates with the morphosyntactic status of the base – a point rightly emphasized by Giegerich (1999). The English adjectival suffix -*able*, for example, behaves as word-level when attached to inflectional stems, but as stem-level when attached to roots or derivational stems: see §5.3.2 below for discussion.
12 Cole (1995: 95) finds no trace of recursive stem-level domains in Spanish, but her assessment is based on the analysis of a single alternation: diphthongization. Bermúdez-Otero (2013a: 67–71, 2016: 408–13) demonstrates stem-level domain recursion in Spanish with evidence from the syllabification of high vocoids.
13 Bermúdez-Otero (2012: 29, 43, 2013b: §36–§37) motivates the existence of analytic listing with psycholinguistic data and with evidence from phrasal idioms, but Köhnlein (2015: 188ff.) shows that analytic listing also provides a solution for a difficult puzzle in the morphophonology of Dutch place names.
14 The first attestation of the verb *accommodate* in the *OED* dates back to 1531; that of *accommodation*, to 1566. An account of stem-level domain recursion driven by nonanalytic listing and morphological blocking generalizes to cases in which the synchronically derived item was borrowed before its synchronic base: Bermúdez-Otero (2012: 37–9, 2013b: §42–§45) shows how, in the course of history, the balance between lexical storage and online grammatical derivation determines whether or not a complex stem-level item develops cyclic behaviour.
15 Wolf (2011: §4) misdescribes Chung's Generalization and cites empirical evidence that has no bearing on its validity. The generalization forbids the cyclic transmission of purely allophonic properties from a base to a stem-level derivative, but freely allows noncontrastive features to be

passed from a base to a word-level derivative. For further clarification of this point with examples from English, see Bermúdez-Otero (2013b: p. 22).

16 As we saw in (17), the explanation of stem-level domain recursion proposed by Bermúdez-Otero (2012, 2013b) holds that, upon first encountering the stem *accòmmodátion-*, speakers redundantly store its stem-level representation in the permanent lexicon. In this view, the representation from which *accòmmodátionlessness* inherits its opaque metrical properties occupies an intermediate position in the static network of lexical relations captured by stem-level redundancy rules, but it does not correspond to an intermediate stage of processing in online speech production. This refinement may be set aside in the current context; the key point is that, in the stratal account, *accòmmodátionlessness* is faithful to the stem-level representation of a stem, and not to the surface form of a word.

17 The superscript $^{\varnothing}$ denotes a lexical gap, i.e. a well-formed but nonexistent lexical item.

18 Steriade claims that stress shift in *remédiable* enables the adjective to satisfy a constraint penalizing word-final strings of three unstressed syllables. This does not quite work, as [ɹi.ˈmiː.djə.bl̩] is only a relaxed variant of the canonical pronunciation [ɹi.ˈmiː.di.ə.bl̩], which has preantepenultimate stress: see Wells (2008: *sub voce*) and note 23 below. In tableau (29), therefore, I substitute a positionally relativized version of Green & Kenstowicz's (1996) foot-based constraint against lapses: this assigns a violation mark for every pair of posttonic unstressed syllables not separated by a foot boundary.

19 I substitute the term *inflectional stem* for Kiparsky's 'word', and *root* or *derivational stem* for Kiparsky's 'stem', in order to maintain consistency with the definitions in (5).

20 In addition, both languages have a very small set of derivational affixes that attach to inflectionally complete grammatical words: see note 8.

21 *Amenable* (< Anglo-Norman *amener* < Latin *mināri*) bears no relation to *amenity* (< Latin *amoenitātem*). The *OED* lists the verb *deléctàte*, with an earliest attestation in 1802 (cf. *c*1400 for *delectable*), but describes it as 'rare' and 'affected or humorous'; it has no tokens in the BNC. The rare verb *dúbĭtàte* has a similar status.

22 This analysis of *-able* provides a straightforward account of the rise of retraction failure in adjectives like *formídable* (cf. earlier *fórmidable*). This can be understood simply as an outcome of univerbation: when an adjective ceases to be parsed as containing the suffix *-able,* it becomes vulnerable to lexically diffusing change towards the general default stress pattern (Bermúdez-Otero 2012: 28). Crucially, this explanation correctly predicts that retraction failure will not occur in novel formations: for an example, see the discussion of **pèriódable* below.

23 As observed in note 18, the relaxed pronunciation [ɹi.ˈmiː.djə.bl̩] derives from the canonical form by a variable low-level process known as 'compression' (Wells 2008: *sub voce*). Compression is a pervasive phenomenon in present-day English. In addition, the vowel length alternation that opposes *remĕdy* to *remēdial*, *remēdiate*, *remēdiation*, and *remēdiable* falls under a general stem-level pattern of CiV-lengthening (Chomsky & Halle 1968: 47; Halle & Mohanan 1985: 78).

24 The following are natural occurrences of *periodable* found online (boldface mine):

(i) Franchisees need to pay a fixed **periodable** fee to franchisors [. . .] http://wiki.answers.com/Q/How_does_a_franchise_operate

(ii) [. . . t]he quotation is a complete **periodable** thought. http://archiver.rootsweb.ancestry.com/th/read/WORDS/2000-01/0947432112

25 Although [ˈpɪə.ɹi.ə.də.bl̩] clearly reflects stress-neutral word-level suffixation, it also happens to be the pronunciation that would arise by stress retraction: cf. e.g. *amélĭŏrable*. Hayes (1982: §2.6.3) accounts for this fact by assuming that the relevant roots have underlying glides: i.e. *per*/j/*od*$_{\surd}$-, -*mel*/j/*or*$_{\surd}$-.

26 Juliet Stanton (personal communication) notes that *périodable* will win in tableau (35) if Ident-σ ⁀Σ° is replaced with some markedness constraint indexed to *-able* and penalizing stress on light antepenultimate syllables. However, there is no independent motivation for a constraint specifically banning stressed light antepenults. Significantly, *-able* did not historically become a weak retractor through a primary innovation, but rather as an opaque by-product of secondary stress reduction: e.g. (*fórmi*)(*dàble*) > (*fórmi*)*d*[ə]*ble*; cf. also (*míli*)(*tàry*) > (*míli*)*t*[ə]*ry* in British English. An analysis of retraction driven by a faithfulness constraint like Ident-σ ⁀Σ° accords well with this fact: an opaquely derived property has been historically reanalyzed as an underlying specification.

27 English adjectives only inflect overtly for degree (e.g. *small-∅ ~ small-er ~ small-est*), but *developmental*, if at all gradable, does so periphrastically (e.g. *more developmental*). In (37,a), however, I have added an inflection node to the structure just to highlight the difference between a stem and a word in the sense of (5).

28 Sproat (1985: 16) and Cole (1995: 87) trace the discussion of bracketing paradoxes in generative linguistics back to Siegel (1974). Other classic works on this topic include Williams (1981), Kiparsky (1983), Pesetsky (1985), Sproat (1985), Hoeksema (1987), and Beard (1991), and the articles by Spencer and Stump cited above.

29 Obligatory stem-level assimilation should not be confused with gradient coarticulation or with the categorical but optional phrase-level process found in some dialects (Bermúdez-Otero & Trousdale 2012: 694–6).

30 *Pace* Newell (this volume), the nasal of the prefix *in-* undergoes place assimilation even across ω-boundaries (Wennerstrom 1993; Raffelsiefen 1999). This is shown by forms like *impolite* ($_{\omega'}$($_{\omega}$ ˌɪm)($_{\omega}$ pʰəˈlaɪt)), where [m] exhibits assimilation but [pʰ] must be foot-initial since it is aspirated; cf. *importune* ($_{\omega}$ ˌɪmpəˈtjuːn). This idiosyncrasy of *in-* is best analyzed as reflecting underspecification: the nasal of /ɪN-/ is thus underlyingly different from that of /ʌn-/. It appears that /ɪN-/ has become a dual-level affix like *-able* (§5.3.2): it displays the prosodic behaviour of a stem-level prefix in root-based forms like *importune*, whereas it is prosodified in the same way as word-level *un-* in stem-based forms like *impolite*.

References

Albright, Adam. 2008. Inflectional paradigms have bases too: Evidence from Yiddish. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity* (Oxford Studies in Theoretical Linguistics 18), 271–312. Oxford: Oxford University Press.

Albright, Adam. 2015. Faithfulness to non-contrastive phonetic properties in Lakhota. Paper presented at the 12th Old World Conference in Phonology, Barcelona, 30 January 2015. Abstract available at www.ub.edu/ocp12/wp-content/uploads/2014/11/OCP12_Main_ALBRIGHT.pdf.

Allen, Margaret Reece. 1978. *Morphological investigations*. Storrs, CT: Doctoral dissertation, University of Connecticut.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21 (4), 737–78.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: The MIT Press.

Aronoff, Mark & S. N. Sridhar. 1983. Morphological levels in English and Kannada; or, *Atarizing Reagan*. In John Richardson, Mitchell Marks & Amy Chukerman (eds.), *Papers from the parasession on the interplay of phonology, morphology, and syntax* (Chicago Linguistic Society 19), 3–16. Chicago: Chicago Linguistic Society.

Artés Cuenca, Eduard. 2016. *The influence of phonology on inflection: The interplay between syllabification and lexical insertion in Pallarese Catalan*. Barcelona: PhD thesis, Universitat Autònoma de Barcelona.

Baker, Brett. 2005. The domain of phonological processes. In Ilana Mushin (ed.), *Proceedings of the 2004 Conference of the Australian Linguistics Society*. Available at http://hdl.handle.net/2123/112.

Beard, Robert. 1991. Decompositional composition: The semantics of scope ambiguities in "bracketing paradoxes". *Natural Language and Linguistic Theory* 9, 195–229.

Benua, Laura. 1995. Identity effects in morphological truncation. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics 18), 77–136. Amherst, MA: GLSA, University of Massachusetts.

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words*. Amherst, MA: Doctoral dissertation, University of Massachusetts. Available at ROA-259–0498, Rutgers Optimality Archive, http://roa.rutgers.edu.

Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Variation within Optimality Theory: Proceedings of the Stockholm* *workshop on 'Variation within Optimality Theory'*, 25–36. Stockholm: Department of Linguistics, Stockholm University. Expanded version available at ROA-593–0403 at the Rutgers Optimality Archive, http://roa.rutgers.edu.

Bermúdez-Otero, Ricardo. 2006. Morphological structure and phonological domains in Spanish denominal derivation. In Fernando Martínez-Gil & Sonia Colina (eds.), *Optimality-theoretic studies in Spanish phonology*, 278–311. Amsterdam: John Benjamins.

Bermúdez-Otero, Ricardo. 2010. Stratal Optimality Theory: An overview. Available at www.bermudez-otero.com/Stratal_Optimality_Theory.htm (accessed on 1 August 2016).

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 4: *Phonological interfaces*, 2019–48. Malden, MA: Wiley-Blackwell.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence* (Oxford Studies in Theoretical Linguistics 41), 8–83. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2013a. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus* 25 (1), 3–103.

Bermúdez-Otero, Ricardo. 2013b. The stem-level syndrome. Paper presented at the Speaker Series of the University of Pennsylvania Linguistics Department, Philadelphia, 11 April 2013. Handout available at www.bermudez-otero.com/stemlevel.pdf.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 374–99. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2016. We do not need structuralist morphemes, but we do need constituent structure. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory* (Linguistics Today 229), 387–430. Amsterdam: John Benjamins.

Bermúdez-Otero, Ricardo. forthcoming. In defence of underlying representations: French adjectival liaison and Romanian morphological palatalization. *Probus*.

Bermúdez-Otero, Ricardo & Ana R. Luís. 2009. Cyclic domains and prosodic spans in the phonology of European Portuguese functional morphs. Paper presented at the Old World Conference in Phonology 6, Edinburgh, 24 January 2009. Available at www.bermudez-otero.com/bermudez-otero&luis.pdf.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 691–720. New York: Oxford University Press.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Bloomfield, Leonard. 1939. Menomini morphophonemics. *Travaux du Cercle Linguistique de Prague* 8, 105–15.

BNC. 2010. The British National Corpus, version 4.3 (BNC*web*, CQP-edition). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Available at www.natcorp.ox.ac.uk/, http://bncweb.lancs.ac.uk/.

Bobaljik, Jonathan David. 2008. Paradigms (optimal and otherwise): A case for skepticism. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity* (Oxford Studies in Theoretical Linguistics 18), 29–54. Oxford: Oxford University Press.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: The MIT Press.

Boersma, Paul & Jan-Willem van Leussen. 2017. Efficient evaluation and learning in multi-level parallel constraint grammars. *Linguistic Inquiry* 48 (3), 349–88.

Booij, Geert. 1997. Non-derivational phonology meets Lexical phonology. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 261–88. Oxford: Clarendon Press.

Booij, Geert & Rochelle Lieber. 1993. On the simultaneity of morphological and prosodic structure. In Sharon Hargus & Ellen M. Kaisse (eds.), *Studies in Lexical phonology*, 23–44. San Diego: Academic Press.

Booij, Geert & Jerzy Rubach. 1984. Morphological and prosodic domains in Lexical Phonology. *Phonology Yearbook* 1, 1–27.

Booij, Geert & Jerzy Rubach. 1987. Postcyclic versus postlexical rules in Lexical Phonology. *Linguistic Inquiry* 18 (1), 1–44.

Borowsky, Toni. 1993. On the Word level. In Sharon Hargus & Ellen M. Kaisse (eds.), *Studies in Lexical Phonology* (Phonetics and Phonology 4), 199–234. San Diego: Academic Press.

Broś, Karolina. 2015. *Survival of the fittest: English and Spanish fricative lenition from the perspective of Optimality Theory*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Buckler, Helen. 2009. *The phonology of word-level suffixes in German and Dutch*. Manchester: MA dissertation, University of Manchester.

Buckler, Helen & Ricardo Bermúdez-Otero. 2012. Word-level affixes trigger stem-level cycles: Evidence from German dorsal fricatives. Paper presented at the 20th Manchester Phonology Meeting, Manchester, 26 May 2012. Available at www.bermudez-otero.com/20mfm.pdf.

Burzio, Luigi. 1996. Surface constraints versus underlying representation. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, vol. 1, 123–41. Salford: European Studies Research Institute, University of Salford.

Bye, Patrik & Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In Jochen Trommer (ed.), *The morphology and phonology of exponence* (Oxford Studies in Theoretical Linguistics 41), 427–95. Oxford: Oxford University Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: The MIT Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Chomsky, Noam, Morris Halle & Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle, Horace G. Lunt, Hugh McLean & Cornelis H. van Schooneveld (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday, 11 October 1956*, 65–80. The Hague: Mouton.

Chung, Sandra. 1983. Transderivational constraints in Chamorro phonology. *Language* 59 (1), 35–66.

CLP, Cercle Linguistique de Prague. 1931. Projet de terminologie phonologique standardisée. In Cercle Linguistique de Prague (ed.), *Réunion phonologique internationale tenue à Prague (18–21/XII 1930)* (Travaux du Cercle Linguistique de Prague 4), 309–23. Prague: Jednota československých matematiků a fysiků.

COCA. 2008. The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. Created by Mark Davies, Brigham Young University, Provo, UT. Available at www.americancorpus.org/.

Cohn, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7 (2), 167–216.

Cohn, Abigail & John J. McCarthy. 1998. Alignment and parallelism in Indonesian phonology. *Working Papers of the Cornell Phonetics Laboratory* 12, 53–137.

Cole, Jennifer S. 1995. The cycle in phonology. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 70–113. Oxford: Blackwell.

Collie, Sarah. 2007. *English stress-preservation and Stratal Optimality Theory*. Edinburgh: PhD thesis, University of Edinburgh. Available at ROA-965–0408, Rutgers Optimality Archive, http://roa.rutgers.edu.

Collie, Sarah. 2008. English stress preservation: The case for "fake cyclicity". *English Language and Linguistics* 12 (3), 505–32.

D'Alessandro, Roberta & Tobias Scheer. 2015. Modular PIC. *Linguistic Inquiry* 46 (4), 593–624.

Deal, Amy Rose & Matthew Wolf. 2017. Outwards-sensitive phonologically-conditioned allomorphy in Nez Perce. In Vera Gribanova & Stepanie S. Shih (eds.), *The morphosyntax-phonology connection*, 29–60. Oxford: Oxford University Press.

De Belder, Marijke, Noam Faust & Nicola Lampitelli. 2014. On a low and high diminutive: Evidence from Italian and Hebrew. In Artemis Alexiadou, Hagit Borer & Florian Schäfer (eds.), *The syntax of roots and the roots of syntax*, 149–63. Oxford: Oxford University Press.

De Vries, J. W. 1984. Adaptation of polymorphemic loanwords: The case of words ending in *-asi* in Indonesian. *Bijdragen tot de Taal-, Land- en Volkenkunde* 140 (4), 476–96.

Embick, David. 2010. *Localism versus globalism in morphology and phonology* (Linguistic Inquiry Monographs 60). Cambridge, MA: The MIT Press.

Embick, David. 2014. Phase cycles, φ-cycles, and phonological (in)activity. In Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi & Nicola Lampitelli (eds.), *The form of structure, the structure of forms: Essays in honor of Jean Lowenstamm*, 271–86. Amsterdam: John Benjamins.

Fabb, Nigel. 1984. *Syntactic affixation*. Cambridge, MA: Doctoral dissertation, MIT.

Fabb, Nigel. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6, 527–39.

Giegerich, Heinz J. 1999. *Lexical strata in English: Morphological causes, phonological effects* (Cambridge Studies in Linguistics 89). Cambridge: Cambridge University Press.

Goedemans, Rob & Ellen van Zanten. 2007. Stress and accent in Indonesian. In V.J. van Heuven & Ellen van Zanten (eds.), *Prosody in Indonesian languages* (LOT Occasional Series 9), 35–62. Utrecht: Landelijke Onderzoekschool Taalkunde.

Green, Thomas & Michael Kenstowicz. 1996. The Lapse constraint. *Proceedings of the 6th Annual Meeting of the Formal Linguistics Society of the Midwest*, 1–15.

Gribanova, Vera. 2015. Exponence and morphosyntactically triggered phonological processes in the Russian verbal complex. *Journal of Linguistics* 51 (3), 519–61.

Hall, T. Alan. 1999. Phonotactics and the prosodic structure of German function words. In T. Alan Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, 99–131. Amsterdam: John Benjamins.

Hall, T. Alan & John H. G. Scott. 2007. Inflectional paradigms have a base: Evidence from s-Dissimilation in Southern German dialects. *Morphology* 17 (1), 151–78.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–76. Cambridge, MA: MIT Press.

Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. In Andrew Carnie & Heidi Harley (eds.), *Papers on phonology and morphology* (MIT Working Papers in Linguistics 21), 275–88. Cambridge, MA: MIT Working Papers in Linguistics.

Halle, Morris & K. P. Mohanan. 1985. Segmental phonology of modern English. *Linguistic Inquiry* 16, 57–116.

Halle, Morris & Jean-Roger Vergnaud. 1987. Stress and the cycle. *Linguistic Inquiry* 18, 45–84.

Hargus, Sharon & Ellen M. Kaisse (eds.). 1993. *Studies in Lexical Phonology* (Phonetics and Phonology 4). San Diego: Academic Press.

Hay, Jennifer. 2003. *Causes and consequences of word structure*. London: Routledge.

Hayes, Bruce. 1982. Extrametricality and English stress. *Linguistic Inquiry* 13 (2), 227–76.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39 (3), 379–440.

Hempel, Carl G. & Paul Oppenheim. 1948. Studies in the logic of explanation. *Philosophy of Science* 15 (2), 135–75.

Hoeksema, Jack. 1987. Relating word structure and logical form. *Linguistic Inquiry* 18, 119–26.

Horn, Laurence R. 2005. An *un*-paper for the unsyntactician. In Salikoko S. Mufwene, Elaine J. Francis & Rebecca S. Wheeler (eds.), *Polymorphous linguistics: Jim McCawley's legacy*, 329–66. Cambridge, MA: The MIT Press.

Hualde, José Ignacio & Marianna Nadeu. 2014. Rhetorical stress in Spanish. In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 228–52. Cambridge: Cambridge University Press.

Inkelas, Sharon. 1989. *Prosodic constituency in the lexicon*. Stanford, CA: Doctoral dissertation, Stanford University. Published (1990), New York: Garland.

Inkelas, Sharon. 1998. The theoretical status of morphologically conditioned phonology: A case study of dominance effects. In Geert Booij & Jaap van Merle (eds.), *Yearbook of morphology 1997*, 121–55. Dordrecht: Kluwer.

Inkelas, Sharon. 2012. The morphology-phonology connection. In Sarah Berson et al. (eds.), *Proceedings of the thirty-fourth annual meeting of the Berkeley Linguistics Society*, 145–62. Berkeley, CA: Berkeley Linguistics Society.

Iosad, Pavel. 2014. The phonology and morphosyntax of Breton mutation. *Lingue e Linguaggio* (1), 23–42.

Iosad, Pavel. 2017. *A substance-free framework for phonology: An analysis of the Breton dialect of Bothoa* (Edinburgh Studies in Theoretical Linguistics). Edinburgh: Edinburgh University Press.

Ito, Junko & Armin Mester. 2009. The extended prosodic word. In Janet Grijzenhout & Barış Kabak (eds.), *Phonological domains: Universals and deviations* (Interface Explorations 16), 135–94. Berlin: Mouton de Gruyter.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51 (3), 639–71.

Jackendoff, Ray. 1997. *The architecture of the language faculty* (Linguistic Inquiry Monograph 28). Cambridge, MA: The MIT Press.

Jakobson, Roman. 1931. Die Betonung und ihre Rolle in der Wort- und Syntagmaphonologie. In Cercle Linguistique de Prague (ed.), *Réunion phonologique internationale tenue à Prague (18–21/XII 1930)* (Travaux du Cercle Linguistique de Prague 4), 164–82. Prague: Jednota československých matematiků a fysiků.

Kager, René. 1989. *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kager, René. 1999. Surface opacity of metrical structure in Optimality Theory. In Ben Hermans & Marc van Oostendorp (eds.), *The derivational residue in phonological Optimality Theory*, 207–45. Amsterdam: John Benjamins.

Kaisse, Ellen M. & Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook* 2, 1–30.

Kaye, Jonathan. 1995. Derivations and interfaces. In Jacques Durand & Francis Katamba (eds.), *Frontiers of phonology: Atoms, structures, derivations*, 289–332. London: Longman.

Kenstowicz, Michael. 1996. Base-identity and uniform exponence: Alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Currents trends in phonology: Models and methods*, vol. 1, 363–93. Salford: European Studies Research Institute, University of Salford.

Kiparsky, Paul. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10 (3), 421–41.

Kiparsky, Paul. 1982a. From Cyclic Phonology to Lexical Phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, vol. 1, 131–75. Dordrecht: Foris.

Kiparsky, Paul. 1982b. Lexical Morphology and Phonology. In In-Seok Yang for the Linguistic Society of Korea (ed.), *Linguistics in the morning calm: Selected papers from SICOL-1981*, vol. 1, 3–91. Seoul: Hanshin Publishing Company.

Kiparsky, Paul. 1983. Word formation and the lexicon. In Francis Ingemann (ed.), *Proceedings of the 1982 Mid-America Linguistics Conference*, 3–29. Lawrence: University of Kansas.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17 (2–4), 351–65.

Kiparsky, Paul. 2003a. Finnish noun inflection. In Diane Nelson & Satu Manninen (eds.), *Generative approaches to Finnic and Saami linguistics*, 109–61. Stanford, CA: CSLI Publications.

Kiparsky, Paul. 2003b. Syllables and moras in Arabic. In Caroline Féry & Ruben van de Vijver (eds.), *The syllable in Optimality Theory*, 147–82. Cambridge: Cambridge University Press.

Kiparsky, Paul. 2005. Paradigm uniformity constraints. Ms, Stanford University. Available at www.stanford.edu/~kiparsky/Papers/LexConservatism.pdf.

Kiparsky, Paul. 2007. Description and explanation: English revisited. Paper presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim, 5 January 2007. Slides available at www.stanford.edu/~kiparsky/Papers/lsa2007.1.pdf.

Kiparsky, Paul. 2010. Reduplication in Stratal OT. In Linda Ann Uyecgu & Lian-Hee Wee (eds.), *Reality exploration and discovery: Pattern interaction in language and life*, 125–42. Stanford: CSLI Publications.

Kiparsky, Paul. 2015a. Phonologization. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 563–82. Oxford: Oxford University Press.

Kiparsky, Paul. 2015b. Stratal OT: A synopsis and FAQs. In Yuchau E. Hsiao & Lian-Hee Wee (eds.), *Capturing phonological shades within and across languages*, 2–44. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kiparsky, Paul. forthcoming. *Paradigms and opacity*. Stanford, CA: CSLI Publications.

Köhnlein, Björn. 2015. The morphological structure of complex place names: The case of Dutch. *Journal of Comparative Germanic Linguistics* 18 (3), 183–215.

Köhnlein, Björn. 2016. Contrastive foot structure in Franconian tone-accent dialects. *Phonology* 33 (1), 87–123.

Lakatos, Imre. 1970. Falsification and the methodology of scientific research programmes. In Imre Lakatos & Alan Musgrave (eds.), *Criticism and the growth of knowledge: Proceedings of the international colloquium in the philosophy of science, London, 1964*, vol. 4, 91–196. Cambridge: Cambridge University Press.

Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8 (2), 249–336.

Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.

Mascaró, Joan. 2016. Morphological exceptions to vowel reduction in Central Catalan and the problem of the missing base. *Catalan Journal of Linguistics* 15, 27–51.

Matzenauer, Carmen Lúcia Barreto & Leda Bisol. 2016. The inventory and the underlying distribution of theme vowels in the Portuguese noun class. *Alfa: Revista de Linguística* 60 (2), 353–65.

McCarthy, John J. 2005. Optimal paradigms. In Laura Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological theory*, 170–210. Oxford: Oxford University Press.

McCarthy, John J. 2007. *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox Publishing.

McCarthy, John J. 2010. An introduction to Harmonic Serialism. *Language and Linguistics Compass* 4 (10), 1001–18.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46 (2), 273–303.

Mohanan, K. P. 1982. *Lexical Phonology*. Cambridge, MA: Doctoral dissertation, MIT.

Mueller, Franz. 2007. Indonesian morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 2, 1207–30. Winona Lake: Eisenbrauns.

Nazarov, Aleksei & Joe Pater. 2017. Learning opacity in Stratal Maximum Entropy Grammar. *Phonology* 34 (2), 299–324.

OED. 1989. *The Oxford English dictionary* (ed. John Simpson and Edmund Weiner), 2nd ed. Oxford: Oxford University Press.

Oostendorp, Marc van. 2012. Stress as a proclitic in Modern Greek. *Lingua* 122 (11), 1165–81.

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology with special attention to Optimality Theory*. Berkeley, CA: Doctoral dissertation, University of California, Berkeley. Available at ROA-171–0197, Rutgers Optimality Archive, http://roa.rutgers.edu.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33, 999–1053.

Pesetsky, David. 1979. Russian morphology and lexical theory. Ms, MIT. Available at http://web.mit.edu/linguistics/www/pesetsky/russmorph.pdf.

Pesetsky, David. 1985. Morphology and logical form. *Linguistic Inquiry* 16, 193–246.

Poser, William. 1989. The metrical foot in Diyari. *Phonology* 6 (1), 117–48.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar* (Technical Reports of the Rutgers University Center for Cognitive Science). New Brunswick, NJ: Rutgers University Center for Cognitive Science. Corrected version (2002) available as ROA-537–0802, Rutgers Optimality Archive, http://roa.rutgers.edu. Published (2004), Oxford: Blackwell.

Raffelsiefen, Renate. 1999. Diagnostics for prosodic words revisited: The case of historically prefixed words in English. In T. Alan Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, 133–201. Amsterdam: John Benjamins.

Raffelsiefen, Renate. 2004. Absolute ill-formedness and other morphophonological effects. *Phonology* 21 (1), 91–142.
Rainer, Franz. 1996. Inflection inside derivation: Evidence from Spanish and Portuguese. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1995*, 83–92. Dordrecht: Kluwer.
Ramsammy, Michael. 2015. The life cycle of phonological processes: Accounting for dialectal micro-typologies. *Language and Linguistics and Language Compass* 9 (1), 33–54.
Roseano, Paolo. 2015. Morfologia non-lineare in romanzo: la flessione verbale del friulano gortano. *Archivio Glottologico Italiano* 100 (1), 85–110.
Rubach, Jerzy. 1984. *Cyclic and lexical phonology: The structure of Polish* (Studies in Generative Grammar 17). Dordrecht: Foris.
Rubach, Jerzy. 1993. *The lexical phonology of Slovak*. Oxford: Oxford University Press.
Saba Kirchner, Jesse. 2010. *Minimal reduplication*. Doctoral dissertation, University of California, Santa Cruz. Available as ROA-1078–0610, Rutgers Optimality Archive, http://roa.rutgers.edu/.
Saba Kirchner, Jesse. 2013. Minimal reduplication and reduplicative exponence. *Morphology* 23 (2), 227–43.
Scheer, Tobias. 2011. *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.
Selkirk, Elisabeth O. 1982. *The syntax of words* (Linguistic Inquiry Monographs 7). Cambridge, MA: The MIT Press.
Shwayder, Kobey. 2015. *Words and subwords: Phonology in a piece-based syntactic morphology*. PhD dissertation, University of Pennsylvania.
Siegel, Dorothy C. 1974. *Topics in English morphology*. Cambridge, MA: Doctoral dissertation, MIT. Published (1979), New York: Garland.
Smith, Peter W., et al. 2016. Case and number suppletion in pronouns. Ms, Goethe-Universität Frankfurt am Main, Syracuse University, and University of Connecticut. Available at http://ling.auf.net/lingbuzz/003110.
Smolensky, Paul. 2006. Computational levels and integrated connectionist/symbolic explanation. In Paul Smolensky & Géraldine Legendre (eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar*, vol. 2: *Linguistic and philosophical implications*, 503–92. Cambridge, MA: The MIT Press.
Spahr, Christopher. 2016. Confronting the European Portuguese central vowel distinction. *Canadian Journal of Linguistics* 61 (2), 211–17.
Spencer, Andrew. 1988. Bracketing paradoxes and the English lexicon. *Language* 64 (4), 663–82.
Sproat, Richard. 1985. *On deriving the lexicon*. Cambridge, MA: Doctoral dissertation, MIT. Available at http://hdl.handle.net/1721.1/15167.
Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison. In J.-Marc Authier, Barbara E. Bullock & Lisa Reid (eds.), *Formal perspectives on Romance linguistics: Selected papers from the 28th Linguistic Symposium on Romance Languages (LSRL XXVIII), University Park, 16–19 April 1998* (Current Issues in Linguistic Theory 185), 243–70. Amsterdam: John Benjamins.
Steriade, Donca. 2008. A pseudo-cyclic effect in Romanian morphophonology. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity* (Oxford Studies in Theoretical Linguistics 18), 313–59. Oxford: Oxford University Press.
Steriade, Donca. 2013. The cycle without containment: Romanian perfect stems. Paper presented at the 21st Manchester Phonology Meeting, Manchester, 25 May 2013.
Steriade, Donca. forthcoming. The cycle without containment: Latin perfects. *Language*.
Stonham, John T. 1994. *Combinatorial morphology* (Current Issues in Linguistic Theory 120). Amsterdam: John Benjamins.
Strycharczuk, Patrycja, Marijn Van 't Veer, Martine Bruil & Kathrin Linke. 2014. Phonetic evidence on phonology-morphosyntax interactions: Sibilant voicing in Quito Spanish. *Journal of Linguistics* 50 (2), 403–52.
Stump, Gregory T. 1991. A paradigm-based theory of morphosemantic mismatches. *Language* 67 (4), 675–725.

Tesar, Bruce. 2014. *Output-driven phonology: Theory and learning* (Cambridge Studies in Linguistics 139). Cambridge: Cambridge University Press.

Trommer, Jochen. 2011. *Phonological aspects of Western Nilotic mutation morphology*. Habilitationsschrift, Institut für Linguistik, Universität Leipzig. Available at www.uni-leipzig.de/~jtrommer/papers/WesternNiloticMutation.pdf.

Trommer, Jochen. 2013. Stress uniformity in Albanian: Morphological arguments for cyclicity. *Linguistic Inquiry* 44 (1), 109–43.

Trommer, Jochen. 2014. Moraic prefixes and suffixes in Anywa. *Lingua* 140, 1–34.

Trommer, Jochen. 2015. Moraic affixes and morphological colors in Dinka. *Linguistic Inquiry* 46 (1), 77–112.

Trommer, Jochen & Eva Zimmermann. 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology* 31 (3), 463–510.

Turton, Danielle. 2016. Synchronic stratum-specific rates of application reflect diachronic change: Morphosyntactic conditioning of variation in English /l/-darkening. *Papers in Historical Phonology* 1, 130–65.

Wallace, Stephen. 1976. *Linguistic and social dimensions of phonological variation in Jakarta Malay*. PhD dissertation, Cornell University.

Wells, J. C. 2008. *Longman pronunciation dictionary*, 3rd edn. Harlow: Longman.

Wennerstrom, Ann. 1993. Focus on the prefix: Evidence for word-internal prosodic words. *Phonology* 10 (2), 309–24.

Wiese, Richard. 1996. *The phonology of German* (The Phonology of the World's Languages). Oxford: Clarendon Press.

Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12 (2), 245–74.

Wolf, Matthew. 2011. Exceptionality. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 4: *Phonological interfaces*, 2538–59. Malden, MA: Wiley-Blackwell.

Zamma, Hideki. 2012. *Patterns and categories of English suffixation and stress placement: A theoretical and quantitative study*. Doctoral dissertation, University of Tsukuba. Available as ROA-1174, Rutgers Optimality Archive, http://roa.rutgers.edu/.

Zdziebko, Sławomir. 2015. A Generalized Nonlinear Affixation approach to Polish palatalizations. *Studies in Polish Linguistics* 10 (1), 17–55.

Zimmermann, Eva. 2013a. Non-concatenative allomorphy is generalized prosodic affixation: The case of Upriver Halkomelem. *Lingua* 134, 1–26.

Zimmermann, Eva. 2013b. Vowel deletion as mora usurpation: The case of Yine. *Phonology* 30 (1), 125–63.

Zimmermann, Eva. 2016a. Copy affixes in Kiranti. In Matías Guzmán Naranjo et al. (eds.), *Replicative processes in grammar* (Linguistische Arbeits Berichte 93), 1–34. Leipzig: Institut für Linguistik, Universität Leipzig.

Zimmermann, Eva. 2016b. The power of a single representation: Morphological tone and allomorphy. *Morphology* 26 (3), 269–94.

Zimmermann, Eva. 2017. *Morphological length and prosodically defective morphemes* (Oxford Studies in Phonology and Phonetics 1). Oxford: Oxford University Press.

6

Rule-based phonology

Background, principles and assumptions

Thomas Purnell

6.1 Introduction

The combined simplicity and elegance of phonological rules attract a great deal of attention across such academic fields as linguistic anthropology, cognitive sciences, linguistics, psychology, sociology etc. *Prima facie*, the rules of human language sounds appear to represent simple accounts of how the world of language sounds works. One reason for phonological rules' broad appeal is their robustness; as "linguistically significant generalizations" (Chomsky & Halle 1968: 330), rules may be stated either in prose or by using formal symbols, and they may be expressed simply or as part of a complex sequence produced by interacting with each other. Phonological rules further function as a gateway leading to morpho-phonological or optimality theoretic analyses (as will be shown elsewhere in this volume). They have also been held up as a serious exemplar in the field of philosophy (Sober 1975).

However attractive phonological rules may be and however useful rules may be to academics of various stripes, since about 1995 rules have been seriously questioned in the field of phonology and have been the locus of a non-amicable field-internal schism. For some contemporary researchers in the field, a phonological rule is an anachronism compared to constraint ranking and thus has little utility outside of teaching undergraduates who are unlikely to continue phonological study. This chapter seeks to remind linguists of the enduring contribution phonological rules have made and continue to make in framing and informing the discussion of cognitive processes involving human language sounds. By reviewing the background, principles and assumptions of the serial application of phonological rules, this chapter highlights the need for clarity in discussions involving distinctive features, rule ordering and interaction with other components (syntax, morphology, semantics and phonetics), given that rules make no sense in the vacuum of formalism alone.

6.2 Historical perspectives

This section presents a prolegomenon to rule formalism, providing some conceptual antecedents necessary to better understand phonological rules, and situates rules within a broader theory of a generative grammar. Formal phonological rules are potentially more powerful than basic descriptive generalizations. Morris Halle has been an influential advocate of phonological rules, arguing in favor of rule formalism and rule ordering, demonstrating how a sequence of sound change processes – which up to that point had been stated most often using prose – might be explained more insightfully with formalism and rule ordering. Halle (1962: 57–58), for example, examines Whitney's *Grammar* (1879), in which four-vowel sandhi processes are described as examples of *coalescing*, *combining*, *converting* and *changing*, shown in (1). Note that although the formalism in the right column of (1) is not present in Halle's paper, it represents a matrix representation of rules he might have proposed at that time. Additionally, the contemporary reader should set aside the urge to write off these examples as an effect of syllabification; review of these rules as a novel presentation of Whitney's descriptions is instructive about the development of rules within the field.

(1) Sanskrit sandhi vowel changes where rules are ordered: (a) > (b) > (c).

	Whitney's 1879 prose (pp. 43ff.)	Halle 1962 style reformulation (pp. 57–58)
a. (=Halle's 5)	"§126. Two similar simple vowels, short or long, **coalesce** and form the corresponding long vowel: thus, two a-vowels (either or both of them short or long) form ā; two i-vowels, ī; two u-vowels, ū; and, theoretically, two ṛ-vowels form ṝ, but it is questionable whether the case ever practically occurs." ati + iva → atī 'va su-uktam → sūktam	$\begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \end{bmatrix}_i + \begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \end{bmatrix}_i \rightarrow \begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \\ +long \end{bmatrix}_i$
b. (=Halle's 7)	"§129. The i-vowels, the u-vowels and the ṛ, before a dissimilar vowel or a diphthong, are regularly **converted** each into its own corresponding semi-vowel, y or v or r." iti + āha → ity āha madhu + iva → madhv iva duhitṛ£-arthe → duhitrarthe	$[+syll] \rightarrow [-syll] / \begin{bmatrix} \_\_ \\ +voc \\ -cons \\ +diffuse \end{bmatrix} + \begin{bmatrix} +syll \\ +voc \\ -cons \\ -diffuse \end{bmatrix}$
c. (=Halle's 6)	"§127. An a-vowel **combines** with a following i-vowel to e; with a u-vowel, to o; with ṛ, to ar; with ḷ (theoretically), to aḷ; with e or āi, to āi; with o or āu, to āu." rāja-indra → rājendra saā + eva → saāi 'va	$\begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \\ +compact \\ -diffuse \\ +grave \end{bmatrix} + \begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \\ -compact \\ +diffuse \\ \alpha grave \end{bmatrix} \rightarrow \begin{bmatrix} \_\_ \\ +syll \\ +voc \\ -cons \\ -compact \\ -diffuse \\ \alpha grave \end{bmatrix}$

d. (=Halle's 8)	"§131. Of a diphthong, the final i- or u-element is **changed** into its corresponding semi-vowel, y or v, before any vowel or diphthong; thus e (really ai . . .) becomes ay, and o (that is au . . .) becomes av; āi becomes āy, and āu becomes āv." ne-a → naya naāi-a → naāya	$[+syll] \rightarrow [-syll] / \left[\begin{array}{c} \\ \overline{} \\ +voc \\ -cons \\ +diffuse \end{array}\right] + \left[\begin{array}{c} +syll \\ +voc \\ -cons \\ -diffuse \end{array}\right]$

NB: the Sanskrit glide <v> approximates the English [w] (§57.a), the Sanskrit <y> approximates the English [j] (§55). The Sanskrit vowel <ṛ> approximates the schwa-r [ɚ] or syllabic [ɹ̩], and the Sanskrit vowel <ḷ> approximates [ʊl], [əl] or [l̩] "in such words as *able*, *angle*, *addle*." (Whitney 1879, p. 11, §24).

Regarding the Sanskrit analysis in (1), Halle's first (implicit) observation was that all of these processes involve a change of one sound to another. The changes in (1) are these: short vowels coalesce to a long vowel (1a); high vowels or schwa devocalize to glides or a rhotic approximant, respectively (1b); low vowels raise or diphthongize (1c); and vowel offglides devocalize (1d). The summary just provided can be simplified even further using such formal devices as matrices of features, changes to the feature matrices, the direction of change (indicated by →) and the triggering environment of the change. In brief, the changes include the following: add the feature [long] (1a); change the valency of [syllabic] (1b) and (1d); and mix valency of [compact], [diffuse] and [grave] of the two adjacent segments (1c).

This direct operationalization of sound modifications leads to Halle's second (explicit) observation that, by serially sequencing the processes, the glide devocalization rule (1d) is eliminated – or more precisely, never activated – because (1d) is a subset to another glide devocalization rule (1b). Crucially, an ordered set of three rules evaluated by the use of formal features provides more insight into the cognitive process of Sanskrit speakers than the four statements in prose. Additionally, this method improves our understanding of Sanskrit. From the mid-twentieth century onward, the hallmark of phonological analyses has been that descriptive processes are stated lucidly using formal notations and that proposals can be straightforwardly evaluated on their ability to account for speech data in as few formal steps as possible. Yet, in spite of the argument by Halle, Chomsky and others that a formal system of rules provides a mechanism for a simple and evaluative explanation of sound systems across languages, the seeds for the mid-1990s schism were set early by the recognition of rule conspiracies and the apparent gradient application of rules, two issues that continue to be of interest.

The first trend of restating rules based on output similarity (rule conspiracy) was a partial return to prose or prose-like explanations of the procedural functions involving human language sounds. Initiated by Kisseberth (1970) and Sommerstein (1977), the point was made that rule formalism was unable to encode more general 'purposes' or functions of rules; this criticism of rules essentially signaled, on the part of some phonologists, a desire to return rules to prose form. Kisseberth's Yawelmani (Yokuts) example shows that rule formalism permits multiple repair strategies – here, Consonant Deletion (2a) and Vowel Insertion (2b) – to apply automatically, blind to the interesting outcome that the two processes "conspire against consonant clustering" (1970: 294). The constraints in Yawelmani are: no clusters at the beginning or end of a word (*#CC, *CC#), and no triconsonantal clusters (*CCC).[1] The rules in (2) reflect structurally similar sub-environments. Kisseberth argues that although these two rules have the same intent or ‘function’ of avoiding complex syllable onsets and codas, the rules have different means of achieving the desired outcome: the deletion rule removes structure while the insertion rule creates structure. An additional issue raised by Kisseberth with these rules is the power of a notational device, specifically braces. This device permits rules to be collapsed into a larger statement of function. Note that with the aid of braces each of the two rules in (2) can be decomposed into two rules, bringing the total number of possible unique operations to four.

(2) Yawelmani conspiracy (Kisseberth 1970: 295–296)

a. Consonant Deletion $C \rightarrow \varnothing \,/\, C \begin{Bmatrix} C + \underline{\quad} \\ + \underline{\quad} C \end{Bmatrix}$ i.e., [CC + C] or [C + CC]

b. Vowel Insertion $\varnothing \rightarrow V \,/\, C \underline{\quad} C \begin{Bmatrix} \# \\ C \end{Bmatrix}$ i.e., [C C #] or [C C C]

In truth, this conspiracy notion was not necessarily a call to return to a Neogrammarian and Americanist bent for descriptions alone, such as the Sanskrit processes stated by Whitney in (1), above. Rather, it suggests that the phonological enterprise should focus on what a process accomplishes, rather than the structure the process applies to. In Kisseberth’s own words:

> The unity of a set of rules may not rest upon the similarity of their structural descriptions, but rather upon the similarity of their *function*. Or to put the point in a slightly different way, rules may be alike in having a common *effect* rather than in operating upon the same class of segments, or performing the same structural change, etc.
>
> *(1970: 293, emphasis Kisseberth’s)*

Some linguists find this shift from formal symbol evaluation to accomplishment evaluation problematic because it imputes intentionality and volition to the system, and not to the speakers. However, others – Chomsky and Halle, for instance – view the presence of rule concatenation via braces an economy afforded by formal devices.

The second challenge to rule formalism has been the problem of gradient vs. categorical change. Where is the dividing point between phonetics and phonology? With respect to rules, when is a process phonetic or phonological? Is the formalism the same for both domains? Early in the development of formal phonological theory the answer was straightforward: phonology is abstract and phonetics is everything else, sometimes referred to simply as ‘implementation.’ That is, anything that appeared to be gradient was ascribed to ‘phonetics.’ Later phonological grammars were partitioned with greater sensitivity to rule domains. For example, Lexical Phonology (Kiparsky 1982) and Postlexical Phonology (Kaisse 1990) made room for different processes at different levels of implementation. One influential line of argumentation held that historical sound changes were repositioned in the grammar, moving along a more transparent phonetics–phonology continuum accounting for how phonological change occurs via phonetics (Ohala 1974, *et seq.*). The loss of a clear separation between phonetics and phonology – or to put it another way, the loss of the autonomy of phonology – has reflexes in contemporary work (e.g., Blevins 2004; Hayes, Kirchner &

Steriade 2004; Port & Leary 2005; Steriade 1995, 1997). Both conspiracies and gradiency have obvious reflexes in non-rule-based phonological theories, primarily Optimality Theory (Prince and Smolensky 1993; McCarthy and Prince 1993, in particular), and in various other theories and approaches in the remainder of this volume.

6.2.1 Rules in the context of a grammar

For the present paper, phonetics and phonology are broadly defined as the scientific studies of sounds and sound systems, respectively, and operate within a broader set of cognitive modules (Figure 6.1).[2] Phonology itself is not seen as a single monolithic module of language, but the concatenation of multiple modules within a highly modular system (Figure 6.2) emphasizing domain specificity and information encapsulation (Fodor 1983). This massive modularity – to apply a term from Carruthers (2003) stretching across the Denes and Pinson (1963: 4) speech chain – is viewed as organic, where an organic phonological system parallels the type of system observed in biological systems of which language is one part. A hypermodular grammar of phonology multiplies the notion of levels well beyond what is seen in Lexical and Postlexical Phonology proposals. Thus, proponents of a massively modular system view the Ohalaian nondiscrete continuum as a flipbook, where each page of the flipbook holds a unique, discrete digital picture which when moved rapidly appears continuous and in an analog state. In other words, the discrete continuum is an analog illusion.

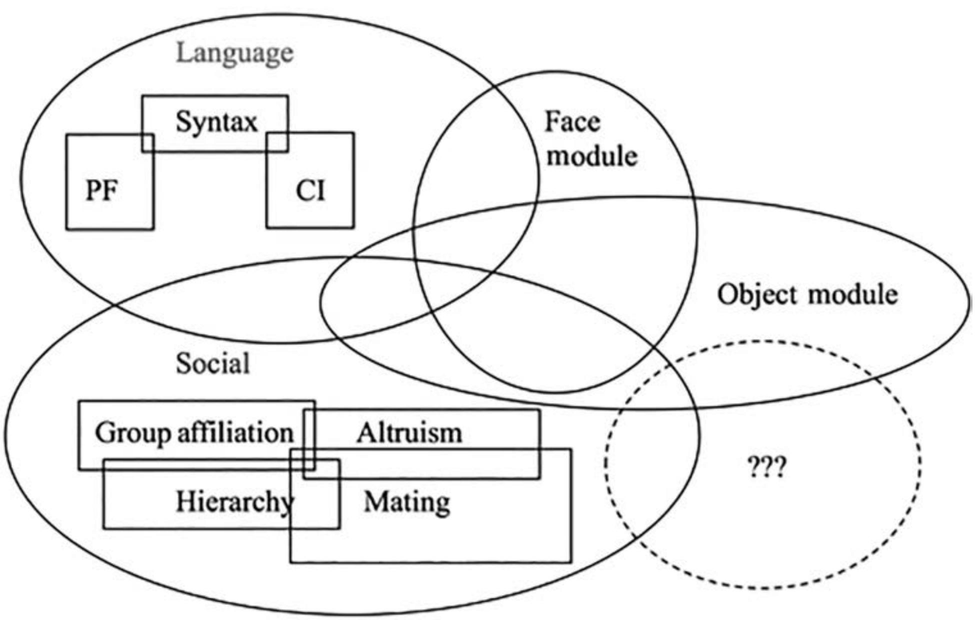


Figure 6.1 Broad picture of linguistic grammar with relations to other cognitive modules (adapted from Purnell, Raimy & Salmons, 2017)

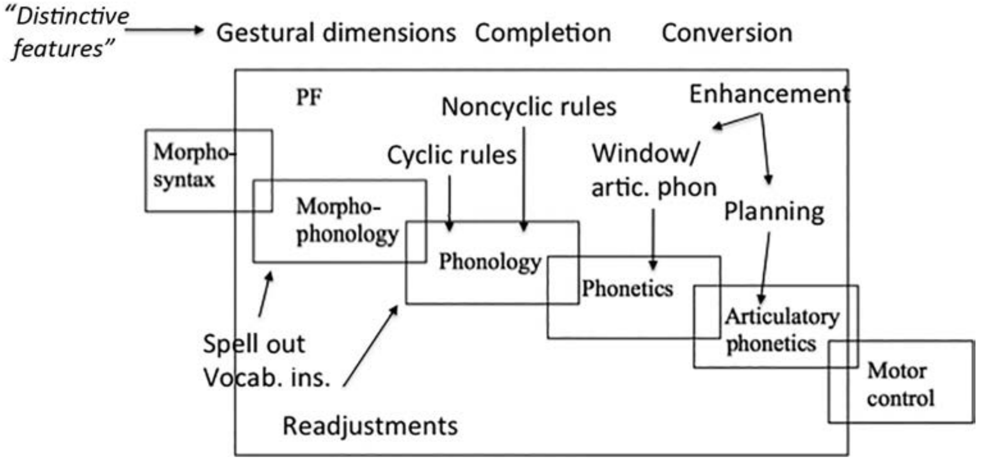


Figure 6.2 Hypermodular grammar of phonology from morpho-syntax to motor control (adapted from Purnell, Raimy & Salmons, 2017)

6.2.2 The roles of rules

The enterprise of formal phonology seeks to model observable, systematic behavior of humans with regard to their speech sounds and to better understand how such behavior arises from largely unseen cognitive functions, historical events and abstract structure. As such, phonology interacts with other parts of grammar and other modules of cognition (Figure 6.1). Rules as algorithms or statements of behavior are central to traditional definitions of formal phonology. Chomsky and Halle define phonology using rules as a central theme: phonology is "the system of *rules* that applies to a surface structure and assigns to it a certain phonetic representation drawn from the universal class provided by general linguistic theory" (1968: 9, emphasis mine). This definition is fairly general and applies widely. While the enterprise of modeling sound systems appears straightforward – thanks in part to the notion of the minimal or near-minimal pair test – the term 'phonology' may be used in two distinct ways, representing different cognitive aspects phonologists must account for (Figure 6.3). These two distinct uses of 'phonology' are as a *regulating* system, and as an *implementation* device. First, phonology has a regulative aspect, by which is meant the rules governing a sound system are sequenced and arranged by 'the phonology.' Phonology, the regulator, is an integral part of any but the most anarchistic theory of human language sounds. Second, phonology has a functional aspect where rules do work. Phonology *qua* function evaluation device then has a role as a grammatical actor. This dual meaning of 'phonology' is similar to the distinction between a procedure (similar to derivation) that is a set of algorithms (the functions) that constitute that procedure (Gallistel & King 2009). To exemplify the interaction of these two aspects of phonology, consider an aspirated apical plosive [t^h] that flaps against the alveolar ridge [ɾ] in many varieties of American English (*butter*, *bottle*, etc.); this process occurs after – and crucially not before – emphasis is increased (i.e., stressed) to the vowel preceding the plosive and emphasis is decreased (i.e., unstressed) on the vowel following the plosive.

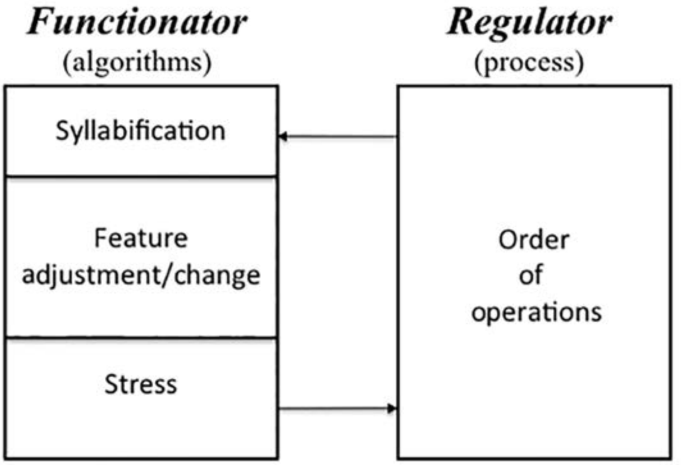


Figure 6.3 Two primary aspects of phonology in a rule-based grammar (adapted from Purnell, Raimy & Salmons, 2017)

In other words, phonology-as-functionator performs the cognitive action of flapping while phonology-as-regulator arranges the flapping action after stress assignment, which itself is ordered after segments have been syllabified into syllables, and so on. The cycling of phonological forms through the functionator is controlled by the regulator.

Any conflation of regulation and function in the literature unfortunately produces imprecise analyses of the behavior of sounds in the speech chain. Much of what Chomsky and Halle (1968) tried to point out was the need for an evaluation metric of both functions and ordering processes. Their evaluation metric is the formal set of phonological notations. Chomsky and Halle begin *The Sound Pattern of English* (1968) with an observation on the relation of phonological rules to a grammar:

> The goal of the descriptive study of a language is the construction of a grammar. We may think of a language as a set of sentences, each with an ideal phonetic form and an associated intrinsic semantic interpretation. The grammar of the language is the system of rules that specifies this sound-meaning correspondence.
>
> *(p. 4)*

Following the pattern set by Chomsky's 'linguistic level' analysis (1965: 11), a generative phonologist is interested in modeling observed speech patterns resulting from sparse, abstract representations held in the mind.[3] Properties and devices of formal phonological models primarily aim to be didactic.[4] That is, first, the model should explain observed speech patterns among the speakers of human languages that are both consistent and variable within and across registers, accents, dialects and languages (Chomsky & Halle 1968: 331). Second, the phonological model should inform more generally the field of cognitive sciences. From ancient scholars (see Cardona 1994; Kiparsky 1995), to the work by Saussure (1959) and the Prague Phonology Circle at the end the nineteenth century, continuing up through the generative tradition (Anderson 1985), and culminating with Archangeli and Pulleyblank (1994) at the end of the twentieth century, generative phonological models consist of three classes of formal devices: objects, functions over the objects and restrictions on those functions,[5] or in other words, representations, rules and constraints, respectively.

Against this background we can understand, in large part, external boundaries of rules. We turn now to examine how contemporary phonological rules are operationalized and how they model cognitive processes involving human language sounds through the characterization of rules as independent functions of objects. We continue with a brief examination of rule formalism. However, a meaningful discussion of rules as functions cannot commence without a brief preliminary description of the class of objects and constraints that a phonological function operates over. While Chomsky and Halle's (1968) rules rely on a feature matrix, and Goldsmith's (1976) Autosegmental Phonological rules rely on autonomous features connected or coindexed to other features, we consider a theory that relies on a combination of Dresher's (2009) contrastive theory and Avery and Idsardi's (2001) distinctive features theory. The parameterization of rules reflecting rule types will be followed by an overview of the types of behavior that rules model and of how rules interact.

6.3 Critical issues and topics

6.3.1 Rule formalism

As noted above, phonological functions – or rules, or algorithms – constitute one of two 'active' portions of the grammar or implicit knowledge of a speaker of a specific language. This active or generative knowledge results in recognizable surface forms allowing communication to proceed. Nevertheless, why subscribe to the notion of rules in the first place? Why not just memorize all the forms in the lexicon? One argument against absolute memorization is that rule systems are computationally efficient. That is, the cognitive cost of absolute storage is greater than the cost to the system balancing storage with algorithms for manipulating stored and derived objects (Gallistel & King 2009). Another reason, and perhaps the most common argument in favor of rules and against absolute memorization – what Kenstowicz and Kisseberth call the "null hypothesis" of phonology (1979: 26ff.) – is that speakers exhibit predictable alternate behavior when confronted with novel forms. For example, the alternation of vocal fold vibration on the plural of novel English words is [z], [s], [z] and [ɪz], depending on the final sound in the novel word. This ability of phonological grammars to predict the correct variable output of the plural alternation is what Gallistel and King (2009) refer to as "combinatorial function." That (a) human language is rife with alternations, (b) the cost of memorizing the entire set of alternations is greater than accessing combinatorial functions, and (c) such alternations work as advertised when novel words are formed or borrowed strongly suggests that the human language faculty taps into a productive algorithm to maintain predictable communicability of speech. It is for these reasons that generative models of phonology have long accepted stating such functions in the form of rules.

Kenstowicz and Kisseberth (K&K; 1979) provide two definitions of a phonological rule, the first one being a broad generalization and the second definition stated more formally:

> The phonological rules, then, make up the phonological component of the grammar, and their function is to convert the [memorized contrasts and idiosyncratic elements] of any utterance into its corresponding [heard pronunciation] by assigning predictable phonetic properties.
>
> *(p. 32)*

> Phonological rules are operations upon strings of feature matrixes. Each rule assigns one or more feature specifications to a matrix when that matrix appears in a certain context. In the statement of the rule, the set of segments undergoing the rule as well as the set of segments which form the context in which the rule operates are identified by mentioning all of the features necessary to uniquely indicate just those particular sets.
>
> *(p. 34)*

The basis of rule formalism is grounded in literature beginning with Halle (1962) and crystalized in Chomsky and Halle's (1968) *Sound Pattern of English* (SPE). SPE (p. 330) lists the two goals ("conditions of adequacy") of a formal theory. First, the formal presentation of rules must be governed by the concept of simplicity (i.e., stated "precisely and clearly" and providing proper "evaluation"; see also pp. 334–5 and the Conciseness Condition, p. 336). Second, the formal presentation of rules should be empirically adequate. Rules should "permit us to formulate general statements about the language which are true and significant (p. 330), and must provide a basis for distinguishing these from other generalizations which are false, or which are true but not significant." The analysis:

> goes beyond [the English] data, as any grammar must, both in depth and in scope – in depth insofar as it expresses the facts that underlie the data, and in scope insofar as it deals with other potential data, with linguistic forms that we did not specifically consider, including indefinitely many that have never been produced.
>
> *(p. 330)*

The formal structure of a phonological rule proposed in SPE is shown in (3) and (4). There are three critical pieces to the formalism in (3a): an input (A), an output (B) and an environment (with X optionally present to the left and Y to the right). In (3a) the underline is where the object A occurs in the description of the rule and where B occurs after the structural change has taken place. The arrow symbol means 'is actualized as' (p. 332; or "rewrite as," Botha 1971: 61), or as a conversion process, as shown in their more explicit definition of a rule (4) which combines both (3a) and (3b).[6] The portion of (3b) to the left of the arrow is termed the Structural Condition that must be present for the rule to apply. The traditional interpretation of this string in (3a) is: 'The phonological object A becomes the object B when X is to A's left and Y is to A's right.' This formal structure is stated and evaluated in either the (3a) form emphasizing the three parts of the rule independently, or as the (3b) form emphasizing the rule's input (the Structural Description of the rule, XAY) and output (the Structural Change of the rule, XBY). The interpretation and effect of (3a) and (3b) are synonymous. If we look back at the rightmost column in the Sanskrit example (1), we observe that devocalization rules (1b,d) are written in the (3a) format, while the coalescence rule making long vowels (1a) and the combination rule (1c) are written in the (3b) format. This alternation in rule format is for explanatory convenience only. At the time of SPE, all of the letters used in a rule's description (A, B, X, Y) represent "feature columns or sequences of feature columns" (Chomsky & Halle 1968: 335), that is, in a feature matrix.

(3) a. A → B/X _ Y
    b. X A Y → X B Y

(4) A rule of the form *A* → *B/X* __ *Y* applies to any string *Z* = . . . *X'A'Y'* . . . , where *X'*, *A'*, *Y'* are not distinct from *X*, *A*, *Y*, respectively; and it converts *Z* to *Z'* = . . .

X'B'Y' . . . , where *B'* contains all specified features of *B* in addition to all features of *A'* not specified in *B* (Chomsky & Halle 1968: 337).

In addition to the objects (A, B, X, Y), the arrow, forward slash and underline, there are additional formal notational devices used in generative phonological rules. The first of these is the curly brace, as seen in the Yawelmani example above in (2), given here as (5). This brace notation ({. . .}) identifies a set of objects so as to collapse multiple environments into one statement. The Vowel Insertion process (5) can be rewritten as two rules of vowel insertion, one applying when two contiguous consonants occur at the right edge of a word (5.i, below) and another when three contiguous consonants occur (5.ii, below) (K&K, pp. 86, 360).

(5)	Vowel Insertion	$Ø \rightarrow V \,/\, C \_\_ C \begin{Bmatrix} \# \\ C \end{Bmatrix}$	
(5.i)	Vowel Insertion 1	Ø → i/C __ C #	/logw-t/ → [logwit]
(5.ii)	Vowel Insertion 2	Ø → i/C __ C C	/lihm + hin/ → [lihimhin]

Chomsky and Halle (1968: 333) point out that representing a rule as a single change to a class of sounds, or schema, stands for the possible individual rules; thus this schema is a more general statement, thus fulfilling the function of a rule in capturing generalizations. Specifically, they state that:

> [t]he characterization of the change as a single process, however, presupposes the existence of rule schemata as entities to which phonological changes may apply. Since schemata exist in a grammar only by virtue of conventions . . . , the examples . . . might be regarded as evidence in support of the reality of rule schemata and the conventions governing their use.
>
> *(1968: 334)*

A further notational device for rules is a mark indicating morpho-syntactic boundaries, when necessary. Various sorts of boundary use various symbols: a hyphen (-), plus sign (+), pound sign (#) and double pound sign (##). Generally, while the hyphen and plus are both morphological boundaries, the hyphen is used most frequently; likewise, most often a single pound sign is used to represent word boundaries (K&K p. 41).

To see the effect of boundaries in rules, consider again the Yawelmani example in (2)/(5). Kisseberth (1970) focuses on a conspiracy of rules. Yet, in later frameworks, e.g., from the perspective of Lexical and Postlexical Phonology onward, the processes described by the rules are situated in different strata, precisely reflecting different boundary marks (morpheme, word and no boundary). Had Kisseberth proposed the rules in the 1980s instead of the 1970s the analysis and conclusions might have been different: we would notice that Consonant Deletion (2a) applies before Vowel Insertion (2b) because of the morpheme boundary, and that one vowel insertion (5.ii) would apply as a late rule because of the lack of boundaries, i.e., after (5.i) which has a word end boundary. The presence of boundary notation would, in effect, weaken the hypothesis that a single conspiracy motivates these processes.

The third additional piece of machinery to consider here is the use of parentheses in the Structural Description of a rule. Parentheses allow for an element to be optionally present in the application of a rule. Consider the two rules in (6) for Karok (Bright 1957, cited in K&K, p. 342) where an /s/ palatalizes to [ʃ] following a high front vowel (/i/) for the 1P-SG form

(6a), even if there is an intervening consonant (6b). Although this process could be written with two separate rules (6a,b), the more concise formalism is to use parentheses and collapse the two rules into just one rule (6c). One caution regarding parentheses is that parenthetical material must be inconsequential to the rule. Returning to the Sanskrit example in (1), the difference between (1b) and (1d) has to do with the presence of [-diffuse] in the triggering environment for (1b) but not (1d); the environments could be written with ([-diffuse]), thereby collapsing the rules.

(6)	a. Palatalization I	s → ʃ/i __	[ʔu-skak], [ni-ʃkak] ‘jump’
	b. Palatalization II	s → ʃ/i C __	[ʔu-ksah], [ni-kʃah] ‘laugh’
	c. Palatalization	s → ʃ/i (C) __	

The fourth notational device is the use of Greek symbols to represent the variable value of a feature. In the Sanskrit rule combining the low with high vowels (1c) we observe the use of a Greek symbol signaling variability in the valiancy of the feature [grave]. Thus, the second element has to have the [grave] feature but the feature could be set for positive ([+grave]) or negative ([-grave]); moreover, because there is an alpha in both the Structural Condition and Structural Change of the rule, it must be the same value. Rather than write out both rules, we can collapse the two rules and use the alpha symbol. In other cases, the alpha can be converted to the opposite setting in the matrix by simply negating the alpha. Thus, in a language that changes a voiceless sound to voiced one and a voiced sound to voiceless one as a dissimilation process, the rule would state [α voice] → [-α voice]. This is read as ‘Whatever the value is for [voice], make it the opposite.’ A final aspect to note is that whatever the setting is for a particular Greek symbol, that setting is represented throughout a given rule. To demonstrate this, let us consider the vowel readjustment rule from SPE (p. 209) applying to certain irregular words that can both front a back vowel (*hold* ~ *held*) and back a front vowel (*tell* ~ *told*). Note that the setting of [α back] in the Structural Description affects the setting for rounding in the Structural Change. The rules in (7a) and (7b) can be collapsed as (7c) using Greek letter notation.

(7) a. $V \rightarrow \begin{bmatrix} -back \\ -round \end{bmatrix} / \begin{bmatrix} \underline{\qquad} \\ +back \end{bmatrix}$

b. $V \rightarrow \begin{bmatrix} +back \\ +round \end{bmatrix} / \begin{bmatrix} \underline{\qquad} \\ -back \end{bmatrix}$

c. $V \rightarrow \begin{bmatrix} -\alpha\, back \\ -\alpha\, round \end{bmatrix} / \begin{bmatrix} \underline{\qquad} \\ \alpha\, back \end{bmatrix}$

Related to parenthetic notation, a final notational device that adds flexibility to rules is the angled bracket notation (<. . .>) in a rule’s Structural Description to sequence the multiple application of a rule. This symbolism signals a type of disjunctive rule relation. Disjunctivity is often seen as a specific rule applying at the expense of a more general rule; angled brackets, however, simply identify the specific rule that applies first, after which then the more general rule (without the angle bracket information) applies. As with the boundary notation, the angled bracket is used infrequently even from the beginning – many have argued that the angled bracket is unnecessary either because rules operate on different cycles or because the Elsewhere Condition (see section 6.2.3) already governs disjunctivity. K&K provide an example of umlaut in Old High German (pp. 354–5), arguing that two separate rules are needed to effect umlauting: one rule (8a and 9a, = primary umlaut, Howell & Salmons 1997) fronts and raises a short [a], while a second rule (8b and 9b, = secondary umlaut, Howell & Salmons 1997) fronts all vowels. Discussion of these rules at this point is intended simply to demonstrate how the two rules might be integrated. (In section 6.3.2 we discuss blocking effects on the specific umlaut (8a, 9a) demonstrating that these rules cannot use the angled brackets but, if we consider them at as synchronic rules, would fall under the Elsewhere Condition.) The effect of (9c) is that the rule can apply iteratively as the environment is met; this would not happen under a simple feeding relationship.

(8) a. Primary umlaut
slagi ~ slegi 'strokes'
gasti ~ gesti 'guests'
b. Secondary umlaut
wurmi ~ wörmi 'worms'
ta:ti ~ tä:ti 'deeds'
noti ~ nöti 'needs'

(9) a. $$\begin{bmatrix} +syll \\ +low \\ +back \\ -long \end{bmatrix} \rightarrow \begin{bmatrix} -back \\ -low \end{bmatrix} \;/\; \_\_\_ \; \mathrm{C}\{\mathrm{i}, \mathrm{j}\}$$

b. $$[+syll] \rightarrow [-back] \;/\; \_\_\_ \; \mathrm{C}\ \{\mathrm{i}, \mathrm{j}\}$$

c. $$\begin{bmatrix} +syll \\ \left\langle \begin{matrix} +low \\ +back \\ -long \end{matrix} \right\rangle \end{bmatrix} \rightarrow \begin{bmatrix} -back \\ \langle -low \rangle \end{bmatrix} \;/\; \_\_\_ \; \mathrm{C}\ \{\mathrm{i}, \mathrm{j}\}$$

6.3.2 Disjunctivity

We have noted several times thus far that there may be some inherent ordering among rules where a more specific rule applies prior to or instead of a more general rule. This disjunctivity demonstrates how rule regulation is accomplished by rule formalism. One instance of disjunctivity suggested above was when a rule included a parenthesis (6) such that the more specific rule effectively blocks the general rule by the presence of an optional C. Another instance was the rule with angled brackets in (9c) where different outcomes resulted. Kiparsky initially proposed the Elsewhere Condition to include two ordering properties: the subset relationship (10a); and that the change (10b) had to be either identical as in (6), slightly off as in (9) or incompatible as in (2). For two rules to be in a subset relation, one rule has to involve less phonological structure than the other rule (e.g., the morpheme boundary in (2a), the consonant in (6b) and the extra features in (9a)). Additionally, the one rule with less structure (i.e., the general rule) must include items that are also present in the rule with more structure. In other words, the more specific rule has the same phonological environment as the general rule plus some extra structure.

(10) Elsewhere Condition (Kiparsky 1973: 94)
Two adjacent rules of the form
$A \rightarrow B/P \_\_\_ Q$
$C \rightarrow D/R \_\_\_ S$
are disjunctively ordered if and only if:
(a) the set of strings that fit PAQ is a subset of the set of strings that fit RCS, and
(b) the structural changes of the two rules are either identical or incompatible.

Upon reconsidering English irregular plural nouns (*oxen* where the plural suffix -*en* blocks the regular -*s*), Kiparsky refined the Elsewhere Condition. The subset principle was reworded as 'properly includes' (11i). Also, the structural changes were argued to be 'distinct' (11ii). The upshot of these changes is that identical rules involving parentheses (6c) are not covered under the Elsewhere Condition.

(11) Elsewhere Condition (Kiparsky 1982: 136–7)
Rules A, B in the same component apply disjunctively to a form Φ *iff*:
(i) the structural description of A (the special rule) properly includes the structural description of B (the general rule), and
(ii) the result of applying A to Φ is distinct from the result of applying B to Φ.
In that case, A is applied first, and if it takes effect, then B is not applied.

The benefit of the Elsewhere Condition is that it is easily evaluated the way any other rule is evaluated, that is, by whether or not the conditions are met, as specified in the rule environment. Consequently, the Elsewhere Condition provides guidance on rule ordering. By way of example let us reconsider Yawelmani Consonant Deletion and Vowel Insertion from (2). Recall that Kisseberth's generalization is that two rules appear to conspire against each other by accomplishing the same thing – breaking up consonant clusters – yet the rules perform what appears to be contradictory actions: deletion and insertion.

When evaluating the symbols of the rules for disjunctivity the rules are best stated in the alternate format, focusing on the Structural Description and Change of the rules. The rules in (2) are restated in (12) and (13). With the additional assumption that Yawelmani grammar builds syllables no more complex than CVC, the rule in both (5.i) and (5.ii) can be collapsed to state that a vowel is inserted prior to any unsyllabified consonant (13). Consonant Deletion, the special rule by virtue of having more symbols, is presented as two rules with subscripts on the consonants. We observe that, although an unsyllabified consonant is a critical part of the Structural Description of the rules, it is not always the unsyllabified consonant (C^*) that gets deleted. Rather, the consonant after the juncture is deleted, leaving a different pair of consonants to be resyllabified as part of the adjacent syllables. Vowel Insertion, the general rule, is properly included in both forms of Consonant Deletion, and the rule consequences are distinct. Thus, the Elsewhere Condition arranges rules so that either form of (12) applies at the expense of (13), and (13) will only apply when both rules in (12) fail to apply.

(12) Consonant Deletion
a. $C_1\ C^*_2 + C_3 \rightarrow C_1\ C_2$
b. $C_1 + C^*_2\ C_3 \rightarrow C_1\ C_3$

(13) Vowel Insertion
$C^* \rightarrow VC^*$

6.3.3 Rule action

The most significant post-SPE change to rule formalism provided a new way of modeling rules with the upshot of addressing the representational objects that rules operate over. Rule-based formalism beginning with SPE employed discrete distinctive feature matrices where features were arranged as we have seen in examples so far, in sets, stacks or unordered features. Regardless of whether one employs feature matrices or other feature representations as objects of rules, distinctive features are considered a class of finite, abstract cognitive objects corresponding to a realization in speech by articulators (Jakobson, Fant & Halle 1951; Sagey 1986, *et seq.*). Early work, such as in SPE, claim that each segment consists of a matrix or conglomeration of featural gestures, along the lines of the examples in (1). The standard gestures for quite some time were those proposed in Jakobson and Halle (1956) and Jakobson, Fant and Halle (1951). In the 1970s, matrix structure gave way to a number of proposals (e.g., Leben 1973; Goldsmith 1976), following the argument that distinctive features were partially independent from each other. Features, while independent, are combined or anchored to a timing tier that represents the 'segment.' The classic example of the independence of a given feature comes from suprasegmental tone stability in Lomongo (14). Goldsmith demonstrated that the tone features in Lomongo were stable, that is, they were not deleted along with other segmental features. In these examples below, although some vowels delete, the tone persists. Note that, although the matrix in (15b) requires linearly ordered sub-matrices to make a rising, full or complex high-low-high pattern, the subsets need to be ordered.

(14) Lomongo (Lovins 1971; cited in Goldsmith 1976: 58ff.)

Before elision	*After elision*	*gloss*
bàlońgó bǎkáé	bàlóngá ˇkáé	'his book'
bánà bǎmǒ	bánǎmǒ	'other children'
bọ̌mọ̌ bòtámbá	bọ̌m̀ộ támbá	'another tree'
bǎtswá là èmí	bǎtswêmí	'you who lead me away'

(15) Partial matrix representation for vowels across the juncture in Lomongo

a. *bàlóngá ˇkáé*, 'his book'

b.

$$\begin{array}{ccc} /o/ & \# & /a/ \\ | & & | \\ \begin{bmatrix} V \\ +high \\ -low \end{bmatrix} & & \begin{bmatrix} & V & \\ (a) & & (b) \\ \begin{bmatrix} -high \\ +low \end{bmatrix} & & \begin{bmatrix} +high \\ -low \end{bmatrix} \end{bmatrix} \end{array}$$

c.

$$\begin{bmatrix} & V & \\ (z) & (a) & (b) \\ \begin{bmatrix} +high \\ -low \end{bmatrix} & \begin{bmatrix} -high \\ +low \end{bmatrix} & \begin{bmatrix} +high \\ -low \end{bmatrix} \end{bmatrix}$$

Goldsmith (1976: 59) states the problem for matrix representations as follows: “how can it be that a tone refuses to be deleted when its vowel is deleted?” For a feature matrix with a tone, there need to be several emendations to the matrix, such as adding an (a) and (b) condition to get a rising or falling tone, or (c) to get a fall-rise pattern preserving the stable tones. However, in spite of changes to the matrix, tone stability is not predicted in a segment-as-matrix proposal because the feature for tones is contained within the matrix, and if the timing tier and the matrix are deleted, then so, too, should the tone – but it is not. The problem is that the matrix formalism presents the *segment* as a single, unified unit – and you cannot delete the segment without deleting the entire unit (matrix of features). Goldsmith’s analysis allows a rule to operate on an autonomous *feature* and not on a *segment.*

The inability of the matrix formalism to account for tone stability led to various arguments for changes in rule formalism. Specifically, it moved the enterprise of phonology towards planar descriptions of spreading features and well-formedness conditions (associating tones to timing tiers and timing tiers to tones; no crossing association lines, etc.). Consider in this regard how a matrix representation has two problems: first in representing features spanning multiple timing tiers: and second, in representing how to spread a tone from one segment to another. In Ngizm (Schuh 1971; Lieber 1987; Peng 1992), high tones only spread to (or are shared with) the next vowel to the right. Goldsmith pointed out two ways to spread tones using matrices: a copying or feature-changing rule (16b.i), or an insertion rule (16b.ii). The proposed autonomous tier representation allows each feature to be on its own tier, and a ‘spreading’ rule spans from segment to segment. Thus rules could be written more easily as (16c). This kind of new representation led in turn to a greater facility in modeling different types of actions, such as spreading and delinking, association, dissociation, etc.

(16) Ngizim spreading

a. /ná-nʊm-ú/ [nánʊ́mꜝú] ‘I constructed’

b. i. $\begin{bmatrix} V \\ +high \\ -low \end{bmatrix}\begin{bmatrix} V \\ -high \\ +low \end{bmatrix} \rightarrow \begin{bmatrix} V \\ +high \\ -low \end{bmatrix}\begin{bmatrix} V \\ +high \\ -low \end{bmatrix}$

b. ii. $\emptyset \rightarrow \begin{bmatrix} V \\ +high \\ -low \end{bmatrix} / \begin{bmatrix} V \\ +high \\ -low \end{bmatrix}$ ___

c. timing tier x x

feature tier H

where ‘H’ refers to the tone feature such as Glottal Tension:[stiff] or Larynx Height:[raised] creating an elevated pitch

6.3.4 Rule typology operationalized

After nearly two decades of autosegmental analyses, Archangeli and Pulleyblank (A&P; 1994), in their Grounded Phonology, operationalized rules in a way that had not been undertaken to that point, building on assumptions of Autosegmental Phonology. (In the socio-historical context of the field, unfortunately, this work appeared just as Optimality Theory (OT) was gaining influence and attention was shifted away from important questions facing rule-based analyses.) The operationalization built on the autosegmental representation for rules, particularly with respect to features

(F-elements), association lines (including a partner type of relation called paths) and prosodic or organizational anchors. To A&P, "a rule defines a well-formed relation between a particular F-element (Argument) and a class of anchors (Targets)" (p. 293). For example, in (16c) the F-element (Argument) would be the high tone feature and the anchor (Target) the tone bearing unit. Their goal was to establish parameters in the properties of association: allowing only two settings for association lines, free or linked; identifying whether there is a line or not; and stating a list of features (binary distinctive features) that are critical for the rule to operate correctly. Moreover, there are co-occurrence restrictions (implicational conditions, a.k.a. positive and negative constraints) that can be framed as either active or avoided paths. For example, A&P (p. 49) argue that the restriction *[+nasal, -voice] for English ("no voiceless nasals") prevents a path from ever forming between the feature settings [+nasal] and [-voice]. A&P make a distinction between an association line and a path, where a path is a special statement about association lines connecting F-elements paradigmatically (p. 49) and the relationship between the F-element and anchor is unique.

A&P's proposed rules have four parameters, which present a more restricted or narrow statement than rules we have covered so far. The four parameters are Function, Type, Direction and Iteration. Each of these parameters is stated in terms of oppositional settings, for example, the setting for Function is Insert or Delete. The effect of this setting seems fairly straightforward in that either structure is being created or destroyed. For example, the settings for the Yawelamani rules in (2) involve different Functions. The setting for Type is a Path (established association lines) or an F-element (a distinctive feature that deletes or is inserted). For a metrical or prosodic rule, the F-element can be something other than a distinctive feature. Putting Function and Type together can delink an association line (Delete + Path) leaving behind a floating feature, or delete a feature (Delete + F-element) that essentially ends up deleting the association line as well. The setting for Direction is either Left-to-Right or Right-to-Left, and for Iteration the setting is straightforwardly either Iterative or Noniterative. Inserting a Path from a high tone to each tone bearing units to the right of the initial association line (like the spreading rule in 16c) would involve the following settings: Function{Insert}, Type{Path}, Direction{Left-to-Right}, and Iteration{Iterative}. The benefit of A&P's work is that we are provided with parameters for understanding rule types. There are limitations with the A&P system given that they assume prosodic structure that collapses syllabic and metrical structure. The analyses that A&P examine are largely featural. However, Idsardi (1992) provides a highly operationalized set of rules for computing stress which can extend A&P's analysis in this domain.

One other mechanism relevant for A&P, which will be discussed below, is a conditional statement on the class of items that is affected by a rule. For example, if rounding affects nonlow vowels, then the nonlow vowels are not part of the Argument Conditions. Conversely, if only morae are the targets for rules, then only morae will be part of the Target Conditions. Feature-filling rules or feature-changing rules will be the result of specifying T(arget)-Structure. Furthermore, an F-element that is linked or free (pre-existing conditions) is handled by A(rgument)-Structure. Grounding Conditions will be discussed below in the discussion of constraint types (section 6.4.2).

6.4 Current contributions

6.4.1 Objects in rules

As noted above, phonological rules are central to traditional generative definitions of phonology. Chomsky and Halle posit two requirements of phonology: an evaluation procedure and "a system of formal devices for expressing rules and a set of general conditions on how these rules are organized and how they apply" (1968: 331). We turn now to some of these general conditions. In the following section, constraints that have traditionally been assumed to trigger or block the operation of rules are discussed.

Reflecting on the differences in the theory of phonological rules from Halle's *Sound Pattern of Russian* (1959) to the present, the most variable aspect between then and now has been the objects or representations over which rules apply. To date, both rules and pre-OT (i.e., non-violable) constraints methodologically are dependent on representations. Rules depend on representations, as two well-known observations point out: First, Halle (1962, *et seq.*) observed that, to quote K&K (p. 240), "phonological rules characteristically refer to natural classes," where natural classes are groups of segments that share some representational structure. Second, the now classic dictum by John McCarthy supplements Halle's view: "primary emphasis should be placed on studying phonological representations rather than rules. Simply put, if the representations are right, then the rules will follow" (1988: 84). Additional evidence of this dependency of rules on representations comes from the observable shift of rules from the mid-1970s rules-as-statements over fully specified matrix-like objects (Chomsky & Halle 1968) to rules-as-graphs in Autosegmental Phonology (Goldsmith 1976). In this section, phonological representations are understood to include such objects as feature dimensions and distinctive features, the timing tier along with the prosodic and metrical tiers, a means for linking or co-indexing features with other segmental information and constraints on representations and rules.

Distinctive features, feature nodes and prosodic categories are considered to be the primary phonological objects that make up the sounds of human language (Sagey 1986; Archangeli & Pulleyblank 1994; Halle, Vaux & Wolf 2000; Avery & Idsardi 2001). The nexus of these objects is the timing tier, for it is around this tier that all other tiers revolve. All rule association lines (Goldsmith 1976) or paths (Archangeli & Pulleyblank 1994) either end in the timing node or are assumed to pass through the timing tier.

Just how the objects (features) and timing tier are related, however, is of some dispute. Generally speaking, those working within in the Prosodic Hierarchy framework (Selkirk 1984) assume that there exists a unique hierarchy connecting terminal features to the prosodic category such as the word and phrase so that all terminal distinctive features are linked together in order for the structure to be well formed. Feature nodes or prosodic units act as 'anchors' to features or feature nodes. Alternatively, in the present paper, the assumption is that the segment and accompanying prosodic model consists of a structure with at least three main tiers or planes centered on a timing unit; this segment consists of multidimensional tiers, as shown in Figure 6.4 (see also Halle 1998: 543, Figure 1).

As depicted in Figure 6.4, a phonological segment consists of a timing node, a set of distinctive features, and prosodic and metrical units. Placing the locus of phonological structure at a single point (the X-tier in SPE) allows a centralized relationship that unites features with prosodic structure (the syllable, foot, phonological phrase, etc.) and metrical units (specifically those in the metrical grid of Halle & Vergnaud 1987; Idsardi 1992). The

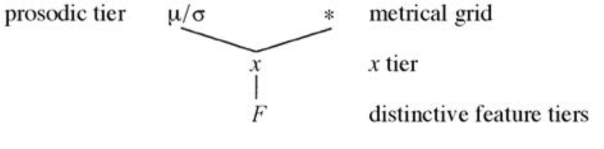


Figure 6.4 Multidimensional representations of a segment at one point in time (from Purnell 1998: 21, Figure 2.2)

central abstract timing unit is a unit of phonological organization that is linearly sequenced with other timing nodes to represent actuation in real time. A phoneme refers to the timing node and the distinctive features associated with it – including only those distinctive features which are critical actors as "distinctive marks on the configurations of words" (Trubetzkoy 1939/1969: 35) necessary for an oppositional contrast. In other words, a phoneme is shorthand for the relation between contrastive distinctive features and the timing tier. The term 'segment' is often used to imply an incomplete or partially underspecified set of distinctive features. When transcribing speech using broad IPA notation or referring to the minimal pair test, for example, it is often assumed that we are explicitly referring to a phoneme. However, the minimal pair test – the first rule for determining phonemes (Trubetzkoy 1939/1969) – is a heuristic device; by contrast, the cognitive comparison between 'phonemes' is more likely a combination of part of the segment (timing node and contrastive dimensions) within a particular context of other segments or segment parts.

With the feature matrix shown to be problematic, the Feature Geometry model (Sagey 1986) remains the most common representational model for features in phonology. Avery and Idsardi (2001), following Gallistel (1980: 58) and Zemlin (1988: 18–27), propose a universal set of features, shown in Figure 6.5, that are bi-oppositional at the statement or feature level (i.e., there are exactly two features per dimension such as [front] and [back]) and organized under articulator nodes, otherwise known as dimensions. Much of the contrast will be stated at these articulator or dimensional levels rather than the feature level. Dimensions, or organizers of the distinctive feature pair, are 'completed' with one of the features in the pair. For example, consider English that has a laryngeal contrast in obstruents. The contrast and rules higher in the cognitive module of formal phonology operate on the Glottal Width dimension. As such, the Avery and Idsardi model differs from other models where the terminal feature spreads (e.g., Revised Articulator Model, Halle, Vaux & Wolf 2000). In the Avery and Idsardi model, only after processes apply are the dimensions completed with actual distinctive features. Avery and Idsardi call this delayed representative stage *completion*.

Purnell and Raimy (2015) combine this notion that contrasts occur at the dimension level of Avery and Idsardi's model with Dresher's (2009) Successive Division Algorithm (SDA),

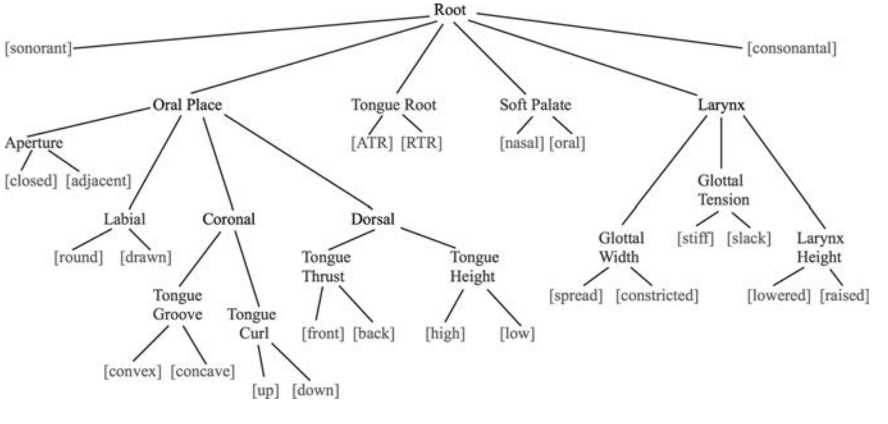
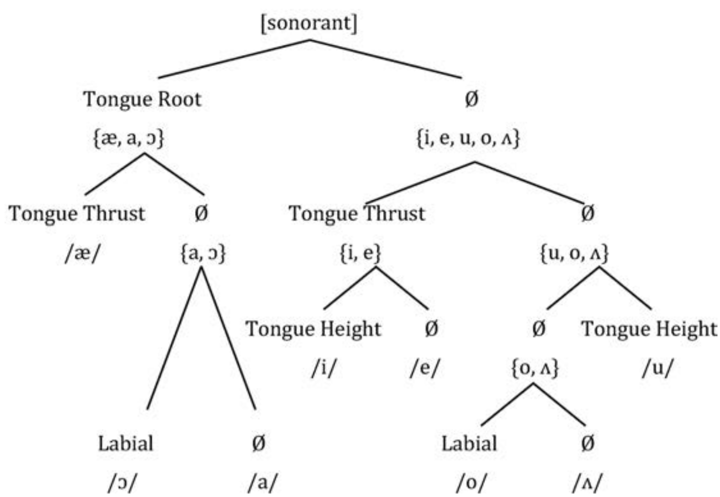


Figure 6.5 Distinctive feature hierarchy depicted by organizational nodes (e.g., Oral Place), dimensions (e.g., Labial) and features (e.g., [round]) that complete the dimensions

(revised from Avery & Idsardi 2001: 66)

which is a language acquisition algorithm, successively distinguishing segments from each other based on one contrastive feature at a time, reducing the phonemes included in a given set with each iteration. Just as the property of representational underspecification was influential at the end of the twentieth century, the SDA also provides an impactful, rule-based analysis. In (17), the top of the feature hierarchy is Tongue Root, separating the low vowels from other vowels. The division continues at lower levels of the hierarchy (17b) until all the sounds have been properly distinguished from other features (17c). One additional change to Dresher's proposal made by Purnell and Raimy is that the contrastive opposition is between dimensions and Ø, not a positive (e.g., [+low]) and a negative (e.g., [-low]) feature setting. This proposal improves the economy of rule-based analyses and incorporates the power of underspecification. As noted earlier, choice of feature theory is nontrivial for rule-based analyses because representations are paramount to rules.

(17) a. Distinctive features of Modern American English long vowels (Purnell & Raimy 2015: 533)



b. Tongue Root > Tongue Thrust > Tongue Height > Labial

c. Modern American English vowel feature specification

Dimension	Completion	I	ɪ	e	ɛ	æ	u	ʊ	o	ɔ	a	ʌ	ə
Tongue Root	[RTR]					√				√	√		
Tongue Thrust	[front]	√	√	√	√	√							
Tongue Height	[high]	√	√				√	√					
Labial	[round]								√	√			
Length	N/A	XX	X	XX	X	XX	XX	X	XX	XX	XX	XX	X

As noted above, features complete the dimensional specification. They do so in two ways. First, the features are made active, filled in etc. on a language specific basis, that is, by rule. This has descriptive implications. For example the hierarchy in (17) informs us as to why there is widespread variation in the pronunciation of American English /u/ but not /i/ (see Labov, Ash & Boberg 2006: 103, Map 10.26). The vowel /u/ across most varieties of American English is marked for Tongue Height alone and the feature [high] completes the Tongue Height dimension by an assignment process (i.e., by rule 18b).

```
(18)  a.   /u/   =        x              (timing tier)
                          |
                      Oral Place
                          |
                      Dorsal Place
                          |
                          TH             (Tongue Height dimension)

      b.    x    →       x          or        Ø → [high] / ___
            |            |                                  |
            TH           TH                                 TH
                         |
                       [high]
```

Also via the hierarchy in (17), the American English vowel /i/ is marked for Tongue Thrust as well as Tongue Height (19a) and is completed by [high] (18b) and [front] (19b).

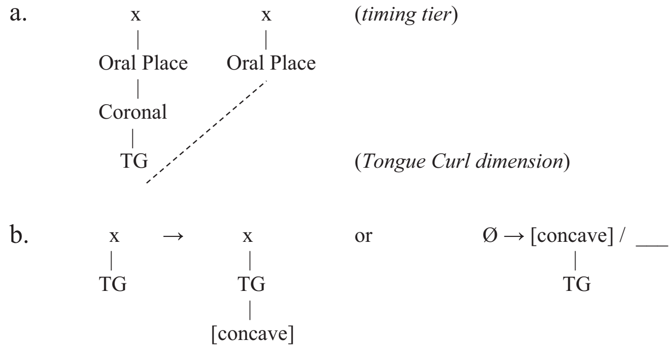
```
(19)  a.   /i/   =        x              (timing tier)
                          |
                      Oral Place
                          |
                      Dorsal Place
                         / \
                       TT   TH           (Tongue Thrust and Height dimensions)

      b.    x    →       x          or        Ø → [front] / ___
            |            |                                   |
            TT           TT                                  TT
                         |
                       [front]
```

This sparser representation afforded by Avery and Idsardi's model (Figure 6.5) and Dresher's SDA for /u/ spoken in the American Midlands and South dialect regions permits /u/ to 'float' forward in the mouth – that is, the vowel is unconstrained by an overt marking by [front] or [back] distinctive feature and can appear as far forward as [ɪ] or as retracted as [u̠]. Thus, a very important empirical fact is captured by not marking /i/ with [+high, +front] and /u/ with [+high, -front], as is often done. Meanwhile, the feature [front] constrains /i/ to the forward region of the vowel space. If an overt feature setting for [back] such as [+back] or [-front] is specified – depending on which framework you follow – then the vowel /u/ cannot float forward. However, we know that /u/ floats and it is desirable that the theory does not restrict it from doing so. But we do want Tongue Height to keep both /u/ and /i/ elevated in the vowel space. Note that although the Southern Shift (see Labov, Ash & Boberg 2006: 125ff., especially Figure 11.2) backs and lowers the /i/, making it a more central vowel, this repositioning in the vowel space, however, is a different process and not one that results from the absence of a feature. That is, /i/ crucially must be effected by a phonological rule to override the contrast derived by the hierarchy.

Secondly, features can be added to a dimension, enhancing a contrast either in a phonological or phonetic module (Stevens & Keyser 1989). Returning to the example of English /u/, most speakers center the vowel when following a Coronal consonant (apicals; see Labov, Ash & Boberg 2006: 101, Map 10.24) by completing the noncontrastive Tongue Groove dimension with [concave] (20b).

(20) Tongue Groove dimension spreading, e.g., *too* /tʰu/ [tʰʉ] ~ [tʰɨ]



When /u/ is not adjacent to a consonant with a Tongue Curl (Coronal) specification, speakers differ in the use of 'no feature,' allowing the vowel to be produced in any high position including the front position (e.g., *cool* as [kʰɨul]~[kʰuwl]), or completing the Labial dimension with [round] producing a perceptually backed and raised variant (e.g., a New England or Great Lakes /u/ [kʰu̝wl]; Labov, Ash & Boberg 2006: 103).

In sum, a description of distinctive features is central to rules because the descriptions (a) establish dimensions and features as types of phonological objects rules manipulate, (b) hint at how functions of these objects cover different ground in formal phonology (feature completion) and early phonetic modules (variability) and (c) reveal the concept of restrained variability in that phonetic variation respects phonologically specified representations (e.g., /u/ stays high in the mouth in American English).

6.4.2 Constraints on rules

The final piece of the phonological machinery is the constraint, that is, how phonological functions over objects are restricted from applying. Traditionally, a constraint is a statement of inclusion or prohibition describing a well-formed construction for a particular language. Stanley (1967: 432–433) defined a constraint as a "negative condition" on phonotactics or rule application, but this is not the only possible type of constraint used in rule-based phonological analyses. In fact, dimension completion rules of Avery and Idsardi along with Grounding Conditions of Archangeli and Pulleyblank (1994: 169–171) are constraints and can be stated either positively or negatively.

Like rules, statement constraints affect phonological action towards well-formed sound shapes; unlike rules, they are generally not considered to be actions themselves. For example, many languages include a constraint that prohibits voiced obstruents in coda position. These are languages that prohibit a path between [consonantal] and Glottal Tension at the end of a syllable. A language like Thai may have a rule (21) which 'repairs' the offensive structure through the deletion of the obstruent or, alternatively, the language may have a rule of devoicing; in A&P terms, there cannot be a path between syllable final [consonantal] and Glottal Tension.

(21) Constraint and repair rule to a prohibition for vocal fold pulsing in syllable final position

a. $\text{*Glottal Tension}]_{\text{syllable}}$

b. $\text{Glottal Tension} \rightarrow \text{Ø}/\_\_\ ]_{\text{syllable}}$

One constraint type is a *phonology-as-functionator constraint*. For example, the application of a rule can be restricted to a certain level (e.g., applies only to certain stems, +, #, etc.). Such constraints are satisfied by the nonapplication of rules. For example, consider the example from English ablaut affecting certain strong verbs, e.g., *sell ~ sold*, but failing to affect regular verbs, e.g., *yell ~ yelled.* Structure Preservation and the strict cycle condition are constraints on rule application of this type.

Another constraint type is that of *positional constraints*. These constraints are often ‘repaired’ by context-sensitive rules. To see how this works in two varieties of American English, we turn to the process known as final devoicing (in African American English) or final fortition (in Upper Midwestern English). Both of these are distinct processes, but they are often referred to with the same label, ‘word final devoicing,’ and may be thought of as multiple ways to avoid voicing at the end of the word. Let us consider the possibility that each effect is driven by the same constraint. These two types of ‘devoicing’ appear to be phonologically driven with active gestures given the laryngeal systems of two respective speech communities.

Final fortition has been argued to occur in the Upper Midwest (Purnell, Salmons & Tepeli 2005; Purnell et al. 2005). It is found in words where a lenis obstruent follows a vowel (*shoes* [ʃuːs]), a rhoticized vowel (*Packers* [p^{h}æː.k^{h}ɚs]) or coda sonorant (*beers* [piɹs], *bans* [pæns]). If we follow the Laryngeal Realism of Iverson and Salmons (1999, *et seq.*) in arguing that [spread glottis] completes the dimension on English aspirated stops /p^h, t^h, k^h/ (corresponding to the orthographic ‘p,’ ‘t’ and ‘k’) while the Laryngeal node is left underspecified for English plain stops /p, t, k/ (corresponding to the orthographic ‘b,’ ‘d’ and ‘g’), then final devoicing is the addition of the Glottal Width specification at the end of the word (22). The rule in (22) provides the plain plosives with Glottal Width dimension and the segment is realized as ‘devoiced’ when [spread glottis] completes Glottal Width.

(22) Upper Midwestern final fortition

Ø → Glottal Width/___ #
|
Aperture (*dimension for stops and fricatives*)

A further type of constraint imposing restrictions on rules is the *juncture constraint*. These constraints, like the positional constraints, are ‘repaired’ by the rule because they occur at the juncture between two morphemes. These effects are unavoidable problems that arise in morphology. We can see this in French hiatus. Calabrese (2009) provides two juncture constraints for dealing with vowel hiatus: no heterosyllabic adjacent vowels, and no triconsonantal consonant cluster at the beginning of syllables. In (23a), the high vowel at the end of the first morpheme devocalizes (raises in the mouth and changes syllabic position); however, in (23b), the high vowel does not devocalize, due to the constraint against three consonants in a row. Instead a glide is inserted as the Onset to the second syllable.

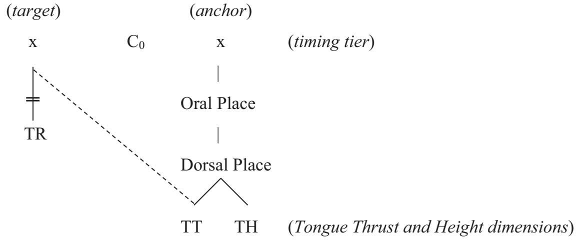
(23) a. /lu+e/ → [lwe]
/li + e/ → [lje]
b. /kri + e/ → [krije], *krje]
/pli + e/ → [plije], *[plje]

It is worth comparing the restriction that Calabrese claims is repaired in (23a) and the statement of devocalization in Sanskrit by Whitney and the rule by Halle. We should ask the question: does the repair occur only because of the presence of the constraint or because of something else (e.g., an identity rule)? Presumably rules operate mechanically and not teleologically (i.e., they don't need a reason to do anything with the right environment). More important is that the structural description of the rule of vowel devocalization includes the violation it seeks to repair. For example, is the point of either the Sanskrit or French vocalization a point about avoiding two vowels in the same syllable that are not a legitimate diphthong, or is it like clusters of vowels, to create a good structure or is it about syllabification? I argue that glide formation is always about syllabification of well-formed syllables, and it is not about avoiding a nonhierarchically represented sequence of sounds.

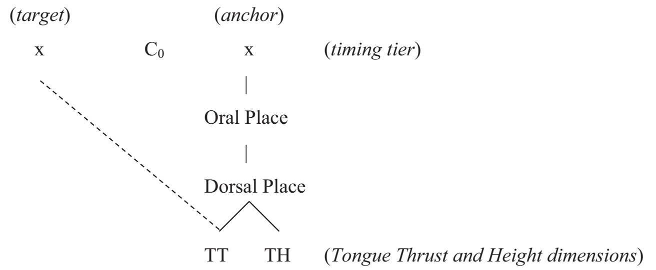
Let us return to the Old High German umlaut example from earlier to demonstrate how a rule-based framework makes use of objects, functions and constraints. Recall that primary umlaut (8a, 9a) appears to be in a disjunctive relation with the secondary umlaut under the Elsewhere Condition; secondary umlaut (8b, 9b) has a subset of formal objects that appear in the primary umlaut; and the two rules do different things by having one rule fronting while the other rule is fronting and raising. For clarity of exposition, one piece of evidence previously omitted was that there is a blocking constraint on the primary umlaut (8a, 9a), specifically that a /l C/ or /x C/ cluster prevents the {i, j} from triggering the fronting and raising of short /a/ (*maht* ~ *mahti* in some dialects, while *maht* ~ *mehti* in other dialects; Howell & Salmons 1997). Howell and Salmons note that while the primary umlaut was an earlier process and the secondary umlaut a later process, there was a time when both were active. Moreover, since the examples above involve plurals, among other morphological forms, the rules would occur on the same stratum.

One possible analysis of the two rules' disjunctivity hinges on feature incompatibility to block spreading. Note that the rule of primary umlaut spreads the F-element Tongue Thrust from an anchor marked with Tongue Height to a preceding target marked only for Tongue Root (24). The reason for raising is that once Tongue Root is lost from the segment and Tongue Thrust is now on the vowel, the vowel is [e], assuming that the vowel system was similar to the English vowel system above, with respect to Tongue Root and Tongue Thrust. The blocking of the rule suggests a co-occurrence constraint prohibiting Tongue Root and Tongue Thrust to be in a single path (e.g., *[Tongue Thrust, Tongue Root]; perhaps a "defective intervention," Nevins 2010: 121–148). One reason, perhaps, that the blocking occurs by an intervening /l/ and /x/ (=[h]) is because they are marked for Tongue Root; Iverson and Salmons (1996) suggest as much, that the short /a/ and /l/ or /x/ share the dimension Tongue Root (see also Howell 1991 for discussion on the unity of the back approximates /l/, /r/ and /x/ in Old English). The marking of Tongue Root prevents a path from making the connection between vowels, stopped by intervening feature specification. Secondary umlaut fails to apply when primary umlaut occurs because it is a more general rule, satisfying the subset and different properties of the Elsewhere Condition in (11). Secondary umlaut (25) is simply a regressive spreading rule that makes no reference to the Tongue Root dimension and hence applies freely across consonants that are marked as such.

(24) Old High German primary umlaut



(25) Old High German secondary umlaut



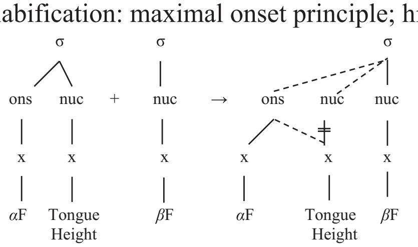
6.4.3 Rule interactions

Central to the rules vs. constraints debate within phonology has been the issue of opacity. Opacity results when the effect of one rule hides the effect of a previous rule, rendering the first rule's output opaque on the surface. Although opacity is dealt with elsewhere in this volume, consider here several types of interactions to demonstrate basic ordering concepts. K&K provide a series of examples that can be restated in a more contemporary rule framework to demonstrate the four ordering types: *feeding* (one rule establishes the environment for the second to take place); *bleeding* (one rule removes the environment, preventing a subsequent rule from applying); *counterfeeding* (a subsequent rule does not take place even though an earlier rule applied); and *counterbleeding* (the trigger for an earlier rule is no longer present even though the rule appears to have applied).

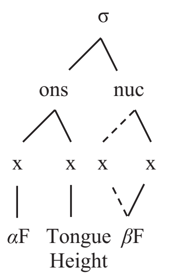
When prefixation results in vowel hiatus in Luganda (Katamba 1989: 123–124), the process of devocalization turns the first of the two vowels into a glide (26a), much like the Sanskrit process in (1). Although described by Katamba as a single process, the rule for vocalization should be accounted for by a sequence of five rules in two general processes: a gliding rule followed by the compensatory lengthening of the vowel features to the open x-slot (based on Hyman & Katamba 1999: 351ff.). Presumably the vowel hiatus requires resyllabification resulting in high vowel gliding when the maximal syllable ONSET is made (27). The adjacent nuclei coalesce to form a single branching syllable nucleus. The distinctive feature dimension of Tongue Height is completed with [high] (28).

(26)	a.	/mu + ana/	[mwa:na]	'child'
		/mu + ojo/	[mwo:jo]	'soul'
		/li + anda/	[lja:nda]	'coal'
		/mi +aka/	[mja:ka]	'years'
	b.	/mu + ti/	[muti]	'tree'
		/mu + kazi/	[mukazi]	'woman'
		/li + no/	[lino]	'this'
		/mi + ti/	[miti]	'trees'

(27) Syllabification: maximal onset principle; high vowel gliding



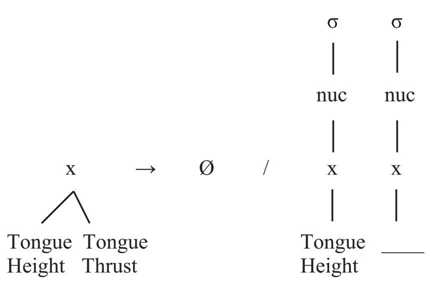
(28) Compensatory lengthening: x-slot insertion; nucleus coalescence; and feature spreading



Katamba (1989) further describes how this process of devocalization interacts with an optional rule of root initial front glide deletion; a palatal glide /j/, when preceded by a prefix of the shape /Cu/, may delete resulting in vowel hiatus (29a). When the prefix is just a single vowel, as in the third person singular prefix in (29b), the glide does not delete. Once front glide deletion takes place (30), then the environment is set up for the devocalization process to take place (i.e., one rule feeds the other).

(29)	a.	/tu + jagala/	[twa:gala] ~ [tujagala]	'we like, we want'
		/ku + jaka/	[kwa:ka] ~ [kujaka]	'to blaze'
		/mu + jola/	[mwo:la] ~ [mujola]	'you-PL carve'
		/tu + jela/	[tw:la] ~ [tujela]	'we sweep'
		/mu + jiko/	[mwi:ko] ~ [mujiko]	'trowel'
	b.	/a + jagala	[ajagala]	's/he likes, wants'
		/e + jaka/	[ejaka]	'it blazes'
		/a + jola/	[ajola]	's/he carves'
		/a + jela/	[ajela]	's/he sweeps'

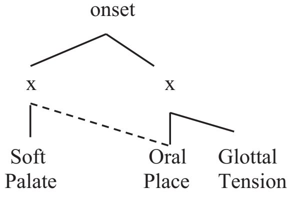
(30) Front glide deletion



To demonstrate how a rule or rules can bleed another rule, we turn to nasals in Swahili. Swahili is a 'voicing' language (Iverson & Salmons 1999), which means that Glottal Tension and Aperture identify fully voiced stops; in contrast, the other category of stops are plain, unaspirated stops (e.g., /p/ not /pʰ/ as in English). Katamba describes a process whereby Glottal Width (completed with [spread glottis]) enhances Aperture (completed with [closed]). That is, when a stop (marked for Aperture) follows a nasal prefix, Glottal Width is added to that segment when Aperture is present and the segment is not specified for Glottal Tension. Once the nasal deletes, then this enhancement of Glottal Width is completed with [spread glottis], hence the aspirated stops on the surface in (31, 32).

(31)	a.	Nasal place assimilation		
		/N + boga/	[mboga]	'vegetable'
		/N + dizi/	[ndizi]	'banana'
		/N + goma/	[ŋgoma]	'drum'
	b.	Nasal deletion and plosive aspiration		
		/N + pange/	[pʰaŋge]	'gadfly'
		/N + taa/	[tʰaː]	'lamp'
		/N + kubwa/	[kʰubwa]	'big-ADJ'

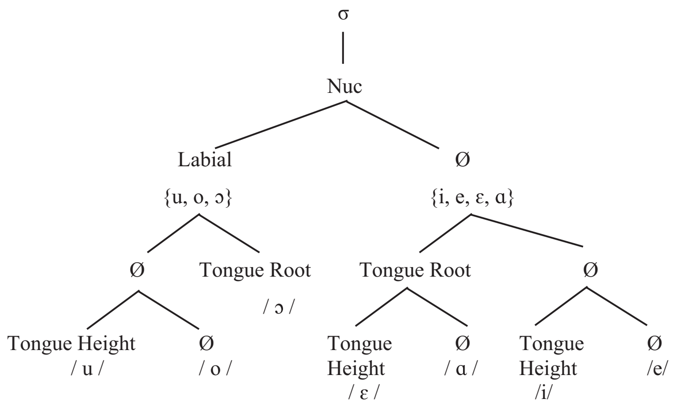
(32) Oral place assimilation



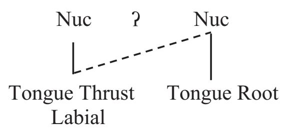
Below is an example from Tunica (Haas 1940, cited in K&K, p. 292ff.) but stated in terms of Dresher-Avery-Idsardi features (34). Data in (33) shows two processes, Vowel Raising (3P-FEM, /-ʔáki/ suffix, 33a) and Syncope (final root vowel in both 3P nonprogressive forms, 33b). Vowel Raising rules are shown in (35) to (37). In the feature method used here, /ɑ/ is the target segment, marked for the Tongue Root. When both Tongue Height and Labial dimensions specified on /u/ spread to the target, /ɔ/ results. However, when only Tongue Height spreads from /i/ to the Tongue Root /ɑ/, [ɛ] results, and when just the Labial dimension on /o/ spreads, /ɔ/ results. The Vowel Syncope, shown in (36), occurs before a glottal stop when the vowel is unstressed (i.e., lacks a line 1 mark on the stress tier). Notice that because of the hierarchy, Tongue Thrust only spreads when present so that /ɑ/ → [ɛ]; otherwise, /ɑ/ → [ɔ]. The contrastive hierarchy simplifies the rules and makes predictions about which vowel quality is selected once the dimensions have been completed with gestures.

(33)		Verb-INF /Ø/	Verb-3P-MASC /-uh-ki/	Verb-3P-FEM /-ʔá-ki/	Verb-3p-Fem-PROG /-hk-ʔɑ-ki/	'gloss'
	a.					
		pó	póʔuhki	póʔɔki	póhkʔáki	'look'
		pí	píʔuhki	píʔɛki	píhkʔáki	'emerge'
		yá	yáʔuhki	yáʔɑki	yáhkʔáki	'do'
		tʃú	tʃúʔuhki	tʃúʔɔki	tʃúhkʔáki	'take'
	b.					
		hárɑ	hárʔuhki	hárʔɑki	hárɑhkʔáki	'sing'
		hípu	hípʔuhki	hípʔɔki	hípuhkʔáki	'dance'
		náʃi	náʃʔuhki	náʃʔɛki	náʃihkʔáki	'lead s.o.'

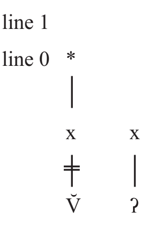
(34) Contrastive hierarchy for Tunica vowels



(35) Tunica Vowel Raising



(36) Tunica Vowel Syncope



(37) Stress rules

a.	Lexical stress:	-ʔáki, line 0 -(x
b.	Line 0 Projection:	all syllable nuclei
c.	Edge Marking:	LLL, #(x
d.	ICC:	n.a.
e.	Repair (Right Destressing):	(→ Ø/(x __
f.	Head Projection:	Left

(38) Derivation example

	/nɑʃi + ʔáki/	‘lead s.o.’
Vowel Raising	nɑʃiʔέki	
Line 1		
Line 0	(x x(x x	
Stress rules (a to c)	náʃiʔέki	
Line 1		
Line 0	(x (x x	
Syncope	náʃʔέki	
Line 1	x	
Line 0	(x x x	
Stress rules (e to f)	náʃʔɛki	

The order of Vowel Raising and Syncope rules exemplifies a counterbleeding order where Syncope removes a portion of the triggering environment for Vowel Raising; if the Vowel Raising and Syncope rules were in reverse order, then Syncope would bleed the Vowel Raising process. Regarding the stress rules, the Right Destressing is just a clash resolution rule. This is important because 'resolution' presumes a prohibition (a.k.a. constraint) against a domain with a single stress-bearing item.

6.5 Future directions

Rule-based analyses continue to successfully model how the human brain uses speech sounds for meaningful human communication. Although the foundation of rule-based analyses continues to be grounded in early work in this line of inquiry, outstanding issues for this approach to phonology reflect both those issues that all theories of phonology face, as well as issues internal to the method of analysis. The pursuit of solutions to most of the issues mentioned here was greatly curtailed by the timing of the redirection away from rule-based analyses. The theory's internal and external issues return us to two challenges for rule formalism noted at the outset of this paper: the role of (disjunctive) 'conspiracies,' and categories versus gradient objects and operations.

The issues all theories and methods of phonological analysis have to attend to are generally at the interface with other domains of investigation. For example, the issue of the relation between phonetics and phonology remains problematic and is at the core of the related interface issues. How does the nature of the functions, objects and boundaries of speech change as information moves towards implementation? The representation of speech described above is largely restricted to more abstract modules. Representations and thus rules in the phonetic domain have different structures. Perhaps these rules are more like windows (Keating 1990) or sliding panes (Browman & Goldstein 1989). The key point in this section is that there are other types of representations and rules in the speech chain, and it is an important challenge for phonological theories to be proposed which respect hypermodularity and features that go along with it, such as enhancement (Stevens & Keyser 1989). Another interface gaining attention by theoretical linguists is variation. Although it was not absent from early works such as SPE, variation has often been viewed as being too messy and unconstrained. However, given that variation within one language cannot exceed the boundaries observed across languages, variation provides a testing ground for parametric studies such as A&P's. Just as discussion on rules stalled out, so too was the discussion in variationist literature the formal representation of variation (the 'variable rule'). The last, and perhaps most important, area of study for phonologists is in the field of cognition. Recent technological advances allow the refinement of information flow through phonological submodules.

Issues inside rule-based approaches need attention as well. These issues include continued testing of A&P's parameterization of rules, streamlining our knowledge of the regulator (e.g., cleaning up the analysis of Halle & Mohanan 1985), and clarifying and maximizing the Elsewhere Condition. The last internal issue that needs work is in the area of formalizing constraints. In this area, more rapprochement with OT is needed: OT is unable to produce a phonology-as-regulator model that is unique from rule-based ones from the 1970s when analyzing opacity and cyclicity effects; and rule-based approaches lack clarity on constraints. Perhaps a way forward would entail discussing whether phonology-as-functionator is structured with functions applying more in parallel while phonology-as-regulator is structured with serial operations.

Further reading

Brain studies inform contemporary phonological research; Poeppel, Emmorey, Hickok and Pylkkänen (2012) overview the flow of information through the brain, suggesting a new model of speech processing. Following forward the research they summarize should lead to a better understanding on cognitive constraints and possibilities for rule-based analyses. Purnell and Raimy (2015) extend Dresher's Successive Division Algorithm in an analysis of Algonquian vowel systems. This work is informative, given the central role representations have with rules, research on underspecification and feature interactions. As noted above, disjunctive rule ordering and the Elsewhere Condition continue to need refinement; this continues to be addressed in works on morphology such as Arregi and Nevins (2013).

Acknowledgements

I am grateful to Eric Raimy, Joe Salmons and Bill Idsardi for decades-long conversations on the nature of rules in phonology and for comments on the paper. All errors are my own and are present to keep those conversations going.

Notes

1 With the help of syllabification, the prohibition in Yawelmani is against unsyllabified consonants at a syllable edge. Vowel Insertion permits incorporating unsyllabified consonants into the word and, thus, the consonant falls in a syllable coda. As a juncture rule, Consonant Deletion has a tangential 'purpose,' targeting the morpheme-initial consonant when an unsyllabified consonant is adjacent to the morpheme boundary.
2 Where *scientific* is further defined as "the intellectual and practical activity encompassing the systematic study of the structure and behavior of [human speech sounds] through observation and experiment" (modified; https://en.oxforddicationaries.com/definition/science, access date 01/10/2017).
3 See also Postal's basic function of phonological theory (1968: 155).
4 A phonological model, unlike a poem, should instruct more than delight, although there is an art – referred to in the field as 'elegance' – to a phonological analysis particularly when rules are involved.
5 Note that the use of the word 'function' throughout should not be interpreted as 'purpose' as employed by Kisseberth (1970) when considering two rules with the same 'function' or effect, for example, repairing syllable structure.
6 The use of the word 'conversion' to characterize rules further distinguishes formal statements from statements in prose such as Whitney's in (1), which also used "coalesce" and "combine."

References

Anderson, Stephen. 1985. *Phonology in the Twentieth Century: Theories of Rules and Representations*. Chicago: University of Chicago Press.

Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.

Arregi, Karlos and Andrew Nevins. 2013. Contextual neutralization and the elsewhere principle. In Ora Matushansky and Alex Marantz (eds.), *Distributed Morphology Today: Morphemes for Morris Halle*, 199–221. Cambridge, MA: MIT Press.

Avery, Peter and William Idsardi. 2001. Laryngeal dimensions, completion and enhancement. In Tracy Alan Hall (ed.), *Distinctive Feature Theory*, 41–70. Berlin: Mouton de Gruyter.

Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Botha, Rudolf. 1971. *Methodological Aspects of Transformational Generative Phonology*. Berlin: Mouton.

Bright, William. 1957. *The Karok Language*. University of California Publications in Linguistics 13. Berkeley, CA: University of California Press.
Browman, Catherine and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6: 201–251.
Calabrese, Andrea. 2009. Markedness theory versus phonological idiosyncrasies in a realistic model of language. In Eric Raimy and Charles Cairns (eds.), *Contemporary Views on Architecture and Representations in Phonology*, 261–304. Cambridge, MA: MIT Press.
Cardona, George. 1994. Indian linguistics. In Giulio Lepschy (ed.), *A History of Linguistics, vol. I: The Eastern Traditions of Linguistics*, 25–60. London: Longman.
Carruthers, Peter. 2003. Moderately Massive Modularity. Royal Institute of Philosophy Supplement, 53, 67–89.
Chomsky, Noam. 1965. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam and Morris Halle. 1968/1991. *The Sound Pattern of English*. Cambridge, MA: MIT Press.
Denes, Peter and Elliot Pinson. 1963. *The Speech Chain*. Murray Hill, NJ: Bell Telephone Laboratories.
de Saussure, Ferdinand. 1959/1966. *Course in General Linguistics*. New York: McGraw-Hill.
Dresher, Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
Fodor, Jerry. 1983. *Modularity of the Mind*. Cambridge, MA: MIT Press.
Gallistel, Charles. 1980. *The Organization of Action: A New Synthesis*. Hillsdale, NJ: Lawrence Erlbaum.
Gallistel, Charles and Adam King. 2009. *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. Oxford: Wiley-Blackwell.
Goldsmith, John. 1976. Autosegmental phonology. Ph.D. dissertation, MIT.
Haas, Mary. 1940. *Tunica, Handbook of American Indian Languages*, Vol. 4. Washington, DC: Smithsonian Institution, Bureau of American Ethnography.
Halle, Morris. 1962. Phonology in generative grammar. *Word* 18: 54–72.
Halle, Morris. 1998. The stress of English words 1968–1998. *Linguistic Inquiry* 29(4): 539–568.
Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale and Samuel Keyser (eds.), *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.
Halle, Morris and K.P. Mohanan. 1985. Segmental phonology of modern English. *Linguistic Inquiry* 16(1): 57–116.
Halle, Morris, Bert Vaux and Andrew Wolf. 2000. On feature spreading and representation of place of articulation. *Linguistic Inquiry* 31(3): 387–444.
Halle, Morris and Jean-Roger Vergnaud 1987. *An essay on stress*. Cambridge, MA: MIT Press.
Hayes, Bruce, Robert Kirchner and Donca Steriade. 2004. *Phonetically Based Phonology*. New York: Cambridge University Press.
Howell, Robert. 1991. *Old English Breaking and Its Germanic Analogues*. Tübingen: Max Niemeyer Verlag.
Howell, Robert and Joseph Salmons. 1997. Umlautless residues in Germanic. *American Journal of Germanic Linguistics & Literatures* 9(1): 83–111.
Hyman, Larry and Francis Katamba. 1999. The syllable in Luganda phonology and morphology. In Harry van der Hulst and Nancy Ritter (eds.), *The Syllable: Views and Facts*, 349–416. Berlin: Mouton de Gruyter.
Idsardi, William. 1992. The computation of prosody. Ph.D. dissertation, MIT.
Idsardi, William and Thomas Purnell. 1997. Metrical tone and the Elsewhere Condition. *Rivista di Linguistica* 9: 129–156.
Iverson, Gregory and Joseph Salmons. 1996. The primacy of primary umlaut. *Beiträge zur Geschichte der deutschen Sprache und Literatur* (PBB) 118: 69–86.
Iverson, Gregory and Joseph Salmons. 1999. Glottal spreading bias in Germanic. *Linguistische Berichte* 178: 135–151.

Jakobson, Roman, Gunnar Fant and Morris Halle. 1951. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge, MA: MIT Press.

Jakobson, Roman and Morris Halle. 1956. *Fundamentals of Language*. The Hague: Mouton.

Kaisse, Ellen. 1990. Toward a typology of postlexical rules. In Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, 127–143. Chicago: University of Chicago Press.

Katamba, Francis. 1989. *An Introduction to Phonology*. London: Longman.

Keating, Patricia. 1990. The window model of coarticulation: Articulatory evidence. In John Kingston and Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 451–470. Cambridge: Cambridge University Press.

Kenstowicz, Michael and Charles Kisseberth. 1979. *Generative Phonology: Description and Theory*. San Diego, CA: Academic Press.

Kiparsky, Paul. 1973. Elsewhere in phonology. In Stephen Anderson and Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 93–106. New York: Holt, Rinehart and Winston.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations, Part I*, 131–175. Dordrecht: Foris.

Kiparsky, Paul. 1995. The phonological basis of sound change. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, 640–670. Oxford: Blackwell.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.

Labov, William, Sharon Ash and Charles Boberg. 2006. *Atlas of North American English*. Berlin: Mouton.

Leben, William. 1973. Suprasegmental phonology. Ph.D. dissertation, MIT.

Lieber, Rochelle. 1987. *An integrated theory of autosegmental processes*. Albany, NY: SUNY Press.

Lovins, Julie. 1971. Melodic conspiracies in Lomongo tonology. In *Papers from CLS 7*, 469–478. Chicago: CLS.

McCarthy, John. 1988. Feature geometry and dependency. *Phonetica* 45: 84–108.

McCarthy, John and Alan Prince. 1993. Prosodic morphology 1: Constraint interaction and satisfaction. Ms., University of Massachusetts and Rutgers University.

Nevins, Andrew. 2010. *Locality in Vowel Harmony*. Cambridge, MA: MIT Press.

Ohala, John. 1974. Phonetic explanation in phonology. In Anthony Bruck, Robert Allen Fox, and Michael W. LaGaly (eds.), *Papers from the Parasession on Natural Phonology*, 251–274. Chicago: University of Chicago, Chicago Linguistic Society.

Peng, Long. 1992. A unified theory of tone-voice. Doctoral dissertation, University of Arizona, Tempe, AZ.

Poeppel, David, Karen Emmorey, Gregory Hickok and Liina Pylkkänen. 2012. Towards a new neurobiology of language. *The Journal of Neuroscience* 32(41): 14125–14131.

Port, Robert and Adam Leary. 2005. Against formal phonology. *Language* 81(4): 927–964.

Postal, Paul. 1968. *Aspects of Phonological Theory*. New York: Harper & Row.

Prince, Alan and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Ms, Rutgers University.

Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. Dordrecht: Reidel.

Purnell, Thomas. 1998. Principles and parameters of phonological rules: Evidence from tone languages. Ph.D. dissertation, University of Delaware.

Purnell, Thomas. 1999. Metrical utility: Extending the power of grids. In Osamu Fujimura, Brian D. Joseph and Bohumil Palek (eds.), *Proceedings of LP'98*, 733–768. Prague: Charles University Press (Karolinum).

Purnell, Thomas. 2009. Phonetic influence on phonological operations. In Cairns and Raimy (eds.), *Contemporary Views on Architecture and Representations in Phonology*, 337–354. Cambridge, MA: MIT Press.

Purnell, Thomas and Eric Raimy. 2015. Distinctive features, levels of representation, and historical phonology. In Patrick Honeybone and Joseph Salmons (eds.), *The Handbook of Historical Phonology*, 522–544. Oxford: Oxford University Press.

Purnell, Thomas, Eric Raimy and Joseph Salmons. 2017. Modularity in phonology. Ms., University of Wisconsin-Madison.

Purnell, Thomas, Joseph Salmons and Dilara Tepeli. 2005. German substrate effects in Wisconsin English: Evidence for final fortition. *American Speech* 80: 135–164.

Purnell, Thomas, Joseph Salmons, Dilara Tepeli and Jennifer Mercer. 2005. Structured heterogeneity and change in laryngeal phonetics: Upper Midwestern final obstruents. *Journal of English Linguistics* 33: 307–338.

Sagey, Elizabeth. 1986. The representation of features and relations in nonlinear phonology. Ph.D. dissertation, MIT.

Sapir, Edward. 1949. The psychological reality of phonemes. In David G. Mandelbaum (ed.), *Selected Writings of Edward Sapir*, 46–60. Berkeley: University of California Press.

Schuh, Russell. 1971. Verb forms and verb aspects in Ngizim. *Journal of African Languages* 10: 47–60.

Selkirk, Elizabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Sober, Elliott. 1975. *Simplicity*. Oxford: Oxford University Press.

Sommerstein, A.H. 1977. *Modern Phonology*. London: Arnold.

Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43(2): 393–436.

Steriade, Donca. 1995. Positional neutralization. Ms., UCLA.

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., UCLA.

Stevens, Kenneth and Samuel Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65(1): 81–106.

Trubetzkoy, N.S. 1939/1969. *Principles of Phonology*. Berkeley, CA: University of California Press.

Whitney, William Dwight. 1879. *A Sanskrit Grammar*. Leipzig: Breitkopf and Härtel.

Williams, Edwin. 1976. Underlying tone in Margi and Igbo. *Linguistic Inquiry* 7: 463–484.

Zemlin, Willard. 1988. *Speech and Hearing Science: Anatomy and Physiology* (Fourth Edition). Boston: Allyn and Bacon.

7
Issues and prospects in Rule-Based Phonology

Bert Vaux and Neil Myler

7.1 Introduction

The term "Rule-Based Phonology" (RBP) has been employed since at least Kaye (1988) as a cover term for phonological theories that derive surface representations from underlying representations via a set of (typically ordered) rules, including not only what McMahon (1992) and others term "Standard Generative Phonology" (Chomsky 1951; Chomsky et al. 1956; Halle 1959, 1962; Chomsky and Halle 1968; Kenstowicz 1994) but also offshoots including Natural Phonology (Stampe 1969, 1973) and Lexical Phonology (Kiparsky 1982, 1985). Since 1993 the term "Rule-Based Phonology" has primarily been employed to represent the standard generative alternative to Optimality Theory, and this is what we concentrate on in this chapter.

RBP shares with Government Phonology (Kaye et al. 1985; Kaye 2000; Scheer 2004, 2012; see also this volume, Chapters 9–11) a rationalist perspective emphasizing formal parsimony and the role of reason in grammar. This perspective sets these theories in opposition to the more empiricist and functionalist orientations of much work in Optimality Theory (OT), Connectionism, Usage-Based Phonology, Probabilistic Phonology, Laboratory Phonology, and Articulatory Phonology.

Work in RBP in the 1960s/70s and the 1980s tended to revolve around the nature of rules and of representations, respectively, but beginning in 1993 the focus of work in RBP shifted away from these to issues raised by OT: Do we need constraints on underlying representations? Is more than one stratum of phonological computation required? Does the phonological component of the grammar employ just rules (Reiss 2008), just constraints (Prince and Smolensky 2003), or both (Calabrese 2005)? Are phonological configurations evaluated locally or globally (Embick 2010)? More generally, as Anderson (2010: 609) observes, "theoretical discussion in phonology since [the] introduction [of descriptions in constraint-based terms] has been largely dominated by comparisons of the two frameworks".

For reasons of space, we lay aside many issues that are of relevance to both RBP and OT but crosscut the two, such as features vs. gestures, feature geometry, alpha notation, the construction and content of underlying representations, the cycle, the Phase Impenetrability Condition, interlanguage effects, and underspecification. In section 7.2, we concentrate on *predictive differences* between rule- and constraint-based formalisms, following David Pesetsky's (2013) aspirational comment that "most of what [linguists] *do* for a living is discover puzzling linguistic phenomena [. . .] and try to explain them by evaluating competing proposals that make distinguishable predictions about the phenomena in question". In section 7.3, we address a debate relevant to both RBP and OT; namely, the question of the content of Universal Grammar (UG). In section 7.4 we discuss future prospects, including a possible cross-over between the rationalist and empiricist perspectives in models using information theory and/or Bayesian probability. Section 7.5 is a brief conclusion.

7.2 Predictive differences between rule- and constraint-based formalisms

When setting out to identify predictive differences between RBP and OT, one encounters at least two potential hurdles. First, some phonologists have suggested that rules and constraints are logically intertranslatable, and more generally that RBP and OT boil down to notational variants of a single generative framework (for variants of this position see Karttunen 1998; Mohanan 2000; and Nevins 2007; see also section 5 of Odden 2008 for discussion of the philosophical and computational issues involved, and McCarthy 1998 for the claim that rules can be thought of as equivalent to a pairing of a markedness constraint with a faithfulness constraint). If rule- and constraint-based formalisms are logically intertranslatable, one might expect there to be no predictive differences between the two. Arguably, though, comparison of *specific instantiations* of RBP or OT (such as the SPE model (Chomsky and Halle 1968) or Classic OT (Kager 1999), respectively) does turn up predictive differences, as we shall see later in this section.

The second hurdle when searching for predictive differences between RBP and OT involves an argument sometimes made by proponents of OT in response to the identification of problematic predictions in specific instantiations of the theory. This argument tends to take something like the following form:

> Scholar X suggests that OT, by virtue of property Y of the theory, incorrectly predicts the impossibility of attested linguistic phenomenon Z. This is not in fact a problem for OT per se, which consists of nothing more than the proposal that grammars are rankings of a universal, violable set of constraints Con.[1] Property Y may be a component of *Scholar A's implementation of* OT, but is not a property of OT.

A stripped-down version of OT of the sort just described makes no testable predictions, though, and hence is not a scientific theory. As Moreton (1999: 5) puts it:

> the claim that every natural-language grammar can be computed by some constraint-hierarchy grammar is [. . .] an empty one. We already know it to be true, since it rules nothing out. [The d]efinition [of a grammar as a ranked set of constraints acting on candidate outputs generated from an input by Gen] provides a framework or notational device, rather than a theory of language. If we want falsifiable predictions, we will have to constrain G[en] and C[on].

Practitioners of OT typically augment Gen and Con with entities such as Eval, Richness of the Base, parallelism, strict domination, Lexicon Optimization, and the rest of what is normally called "Classic OT", as set out in Kager (1999), and it is this more fleshed-out form of OT whose predictions we will be discussing in this chapter.[2] The form of RBP to which it is compared is the one described by Kenstowicz (1994), which assumes that the surface representation of the morphemes in a sequence is derived from their underlying representations by the application of a series of ordered rules. These rules may be subject to the cycle, Structure Preservation, the Derived Environment Condition, and inviolable constraints on underlying and surface representations such as the Obligatory Contour Principle.

Comparison of RBP and OT has thus far been socio-politically inconclusive, in the sense that neither camp has conceded the intellectual high ground to the other. Supporters of OT generally consider conspiracies and the Duplication Problem to pose fatal problems for RBP, while supporters of RBP find these arguments unpersuasive and see insuperable flaws in the OT framework. Some even maintain that the rise of OT has precipitated a decline in the field; as Wauquier et al. (2015) put it,

> the theory that dominated the field since the early 90s, Optimality Theory, rapidly loses velocity and is progressively dissolving. The vacuum that this movement has already created and continues to produce is typically engaged by approaches which have little theoretical substance.

Nevertheless, one can say that the rise of OT as an alternative to RBP has at least had the positive effect of raising issues that were not previously on phonologists' radar, such as overapplication in reduplication (McCarthy and Prince 1999: 290) and Counterfeeding from the Past (Wilson 2006; Wolf 2010).

In this section we examine predictive differences adduced in the literature as arguments in favor of OT (7.2.1) or RBP (7.2.2), focusing for reasons of space on the most-discussed cases, conspiracies and opacity. We try moreover to advance the debate by considering ways in which these arguments have been or might be responded to, and problems that they raise.

7.2.1 Predictive differences suggested to favor OT

Though practitioners of OT do not generally consider RBP as something that needs to be argued with in any depth (Scheer 2010: 195; to appear), when discussing the paradigm shift from RBP to OT they often refer to two predictive differences believed to favor OT over RBP, involving conspiracies and the Duplication Problem respectively. The latter has been addressed in detail by Paster (2013), who argues that the diachronic origins of alternations and lexical patterns explain the frequent cooccurrence of productive alternations and seemingly related static generalizations about underlying forms, such that the relation between the two does not necessarily have to be formally captured in the synchronic grammar. We therefore focus here on conspiracies.

7.2.1.1 Conspiracies

It is commonly argued that OT improves upon RBP by providing an explanatorily superior account of conspiracies. As McCarthy (1999) states, "compelling examples of homogeneity of target/heterogeneity of process tend to support constraint-based over rule-based theories" (cf. Kager 1999: 56; Pater 2001: 161; Hayes 2004: 165). The alleged problem raised by conspiracies for RBP is usually formulated as follows. Kisseberth (1970) noted that in many languages one could observe different rules whose action systematically ensured that a certain output configuration never surfaces. Because rules are independent grammatical processes formulated by the linguist (and, by hypothesis, the Language Acquirer) in accordance with the data, there is no way in RBP (by which he meant SPE, which contained no constraints of the relevant sort) to express the apparent "functional unity" of such rules in the synchronic grammar. In Lombardi's (2002: 13) words, "in a theory where phonological rules specify both context and change, as in SPE and much work following it, it is not possible to account for this asymmetry of patterns except by stipulation".

OT is often argued by its proponents to offer a solution to this problem (see for example McCarthy 2002; Kager 1999). Because grammatical processes in OT result from the interplay of ranked constraints rather than having any independent existence, OT allows for the "functional unity" of conspiratorial processes to be "factored out" (in Kisseberth's 1970 terms) from the processes themselves. This separation of target from repairs is effected as follows:

- The "functional unity" of a set of rules and/or constraints is captured by the fact that the avoided configuration is punished by a single markedness constraint – what McCarthy (2002: 93) terms "homogeneity of target".
- The nature of the repair will depend on (i) the ranking of the other constraints in the grammar – thus cross-linguistic "heterogeneity of repair" (McCarthy 2002: 93) is accounted for; and (ii) the particular environment that a particular marked configuration appears in – some repairs might be good in one situation and less so in another. Hence "heterogeneity of repair" within a language is potentially explicable as well.

*7.2.1.1.1 The putative *NÇ̥ conspiracy*

A famous case study of a conspiracy from the OT literature is Pater's (1999) discussion of the proposed *NÇ̥ constraint in Austronesian languages. Pater begins by observing the case of nasal substitution from the paradigm of the Indonesian prefix *məN-* (on which see also Halle and Clements 1983: 125). Where this prefix is added to root with an initial voiceless stop, the nasal disappears and the stop is replaced by a homorganic nasal (1a). If the root begins with a voiced stop, the nasal assimilates to the place of the stop and there is no deletion (1b). However, combinations of nasals and voiceless stops inside a root are not affected by this process (1c) (Pater 1999: 2, his (1)):

(1)	a.	/məN+pilih/məmilih	'to choose, to vote'
		/ məN+tulis/mənulis	'to write'
		/ məN+kasih/məŋasih	'to give'
	b.	/məN+bəli/məmbəli	'to buy'
		/məN+dapat/məndapat	'to get, to receive'
		/məN+ganti/məŋanti	'to change'
	c.	əmpat 'four' untuk 'for'	muŋkin 'possible'

Pater points out that a variety of processes in Austronesian languages and in the world's languages more generally are similar to nasal substitution in that they eliminate an NÇ̥ cluster – these include denasalization, post-nasal voicing, and nasal deletion. Pater identifies this as a conspiracy to avoid NÇ̥ configurations, and proposes to capture it as follows. The "functional unity" of these processes is explained by positing a *NÇ̥ constraint. A range of faithfulness constraints militate against each of the imaginable repairs for the NÇ̥ configuration, and their ranking determines which repair surfaces in a given language or phonological context.

Nasal substitution (as seen in Indonesian in (1)) is analyzed as Fusion, a process whereby two adjacent segments merge into one. Under the Correspondence theory of Input–Output relations (McCarthy and Prince 1995), Fusion does not violate Max constraints, which punish wholesale deletions. However, it does violate LIN(EARITY), which requires the linear order of output segments to be faithful to the order in the input.[3] Hence, a language where NÇ̥ configurations are eliminated by Fusion can be derived by ranking LIN below *NÇ̥, and by ranking the constraints that punish other repairs above *NÇ̥.

(2) Nasal substitution (cf. Pater 1999: figure (7))

/məN+pilih/	other FAITH	*NÇ̥	LIN
✓ [məmilih]			*
[məmpilih]		*!	

Two other possible repairs for *NÇ̥ are segmental deletion and insertion. The relevant constraints in these cases are DEP and MAX. Beginning with deletions, there is a curious generalization that the nasal is always eliminated in this configuration, never the obstruent. Pater points out (p. 14) that no satisfactory explanation for this is known.[4] The Kelantan dialect of Malay is an example of a language of this sort. The following tableau illustrates how this result is derived (Pater 1999: 15, his (10)):

(3) Tableau for Kelantan-like languages

/NT/	*NÇ̥	OBSMAX	NASMAX
[NT]	*!		
[N]		*!	
✓ [T]			*

It seems that no known language repairs *NÇ̥ violations via epenthesis, a gap that is not predicted by Pater's factorial typology.

A third option is to denasalize the offending nasal. This is what happens in Toba Batak, Kaingang, and Mandar, as illustrated here for the latter (Pater 1999: 16, his (11)):

(4) Mandar maN- prefixation

a. /MaN+dundu/ mandundu 'to drink'
b. /MaN+tunu/ mattunu 'to burn'

Denasalization is effected via deletion of a [nasal] feature.[5] Such feature elimination is punished by IDENT(ITY) constraints. Languages like Mandar can thus be derived by ranking the IDENT[nasal] constraint below *NÇ̥, with all other relevant faithfulness constraints above *NÇ̥ (Pater 1999: 20 his (16)):

(5) Mandar denasalization

/maN-tunu/	DEP	MAX	LIN	*NÇ̥	IDENTNAS
[manunu]			*!		
[mantunu]				*!	
✓[mattunu]					*
[matunu]		*!			
[maŋatunu]	*!				

A fourth option is to voice the obstruent. Pater's illustration of this repair comes from Puyo Pungo Quechua (p. 21, his (17)), where the voicing process occurs only in derived environments[6] (root-internal instances of NÇ̥ are allowed):[7]

(6) a. Root-internal NÇ̥
 šiŋki 'soot' čuntina 'to stir the fire'
 pampal^{y}ina 'skirt'
 b. sinik-pa 'porcupine's' kam-ba 'yours'
 sača-pi 'in the jungle' hatum-bi 'in the big one'
 wasi-ta 'the house (acc.)' wakin-da 'the other (acc.)'

This repair is punished by another Ident constraint, this time Ident[ObsVoice], which requires obstruents in the output to maintain the voicing specification that their input forms have. Thus, the ranking OtherFaith > *NÇ̥ > Ident[ObsVoice] is needed to ensure this result.

All of the constraints so far and the ranking needed to derive nasal substitution are depicted in this tableau (Pater 1999: 23, his (20)):

(7) Pater's final tableau for nasal substitution

/məN-pilih/	Dep	IdentNas	Max	RootLin	IdentObsVoi	*NÇ̥	Lin
✓ [məmilih]							*
[məmpilih]						*	
[məppilih]		*					
[məmbilih]					*		
[məpilih]			*				
[məŋəpilih]	*						

The factorial typology of this system also predicts that there should be languages where one repair is used in the default case, but a different repair is used in situations where the first repair is for some reason blocked. This prediction is seemingly vindicated in certain African languages in which nasal substitution occurs in combinations of nasals and voiceless stops, but nasal deletion is invoked to deal with combinations of nasals and fricatives (p. 24, Pater's (22)):

(8) a. Umbundu (Schadeberg 1982)
 /N+tuma/ [numa] 'I said'
 /N+seva/ [seva] 'I cook'
 b. Si-Luyana (Givón 1970)
 /N+tabi/ [nabi] 'prince'
 /N+supa/ [supa] 'soup'

The constraint ranking *NÇ̥ ≫ Ident [continuant] ≫ Max ≫ Lin derives such behavior (Pater 1999: 25–26, his (23) and (24)):

(9) Ranking for (8)

/N-tabi/	*NC̥	IDENTCONT	MAX	LIN
[ntabi]	*!			
✓[nabi]				*
[tabi]			*!	
/N-supa/				
[nsupa]	*!			
[nupa]		*!		*
✓[supa]			*	

Having introduced the apparent conspiracy between nasal substitution and nasal deletion, Pater argues that his OT approach is superior to a rule-based account on the grounds that:

> [u]nder a purely rule-based analysis [. . .] the functional connection between nasal substitution and nasal deletion would have to be stated independently of the rules themselves; their shared property of eliminating NC̥ clusters is only obliquely retrievable from the rule formulation. This contrasts with the present Optimality Theoretic analysis of African nasal substitution and nasal deletion, in which the functional motivation for these processes is directly incorporated into the formal explanation, thus allowing for a perspicuous account of the conspiracy between them.
> *(1999: 26)*

This claim to superiority seems under-motivated to us, insofar as this factorial typology also yields an array of problematic predictions about intra-language variation when paired with OT devices for capturing optionality. (See also Blust 2004 for a catalog of empirical and conceptual problems with Pater's argumentation.) We now show this by looking at the Constraint Tie approach. Similar problems can be shown to arise for other approaches to optionality in OT, including the Gradual Learning Algorithm (Boersma 1999; Boersma and Hayes 2001) and Markedness Suppression (Kaplan 2011), but we cannot discuss these here for reasons of space.

Constraint ties are a standard device for capturing free variation in the OT literature (see for example McCarthy and Prince 1995: fn 59). Tied constraints are said to be evaluated in all of their logically possible rankings, so that the grammar will be able to produce more than one grammatical output for a single underlying form.

This mechanism has bizarre consequences when it interacts with Pater's factorial typology. Firstly, it leaves open the possibility of a language with complete freedom of choice between two repairs. For example, the repair for a *NC̥ violation in Mandar is to denasalize the nasal consonant. Its near neighbor Konjo, however, instead nasalizes the obstruent and creates a geminate nasal. Pater describes how these two language types can be derived in his system as follows:

> [to derive Konjo], IDENTO→I[NAS] can be ranked beneath IDENTI→O[NAS], so that having an Output nasal in correspondence with an Input obstruent (i.e. NT→NN) is a better resolution of *NC̥ than having an Input nasal in correspondence with an Output obstruent (i.e. NT→TT). In Mandar, of course, the ranking between these constraints would be reversed.
> *(1999: 20)*

The OT mechanism for constraint ties yields a third option: a mixed language in which IDENTO→I[NAS] and IDENTI→O[NAS] are tied. In such a language, both *mattunu* and *mannunu* would be grammatical outputs for /maN+tunu/, and the same optionality would reproduce itself in all combinations of a nasal and a voiceless obstruent in the language. In effect, OT allows for what we call *Conspiratorial Cascades*, an apparently unattested situation in which each individual manifestation of a conspiracy is optionally repaired by the same set of two or more tied phonological operations.

(10) Tied F constraints → tied conspiracies

/maN-tunu/	DEP	MAX	LIN	*NC̥	IDENTI-ONAS	IDENTO-INAS
[manunu]			*!			
[mantunu]				*!		
✓[mattunu]					*	
✓[mannunu]						*
[matunu]		*!				
[maŋatunu]	*!					

The same problem is replicable for any given pair of repairs, even such unlikely bedfellows as epenthesis and deletion. We are not aware of any such cases attested in the literature. Rule-based analyses do not make analogous predictions.

7.2.1.1.2 Conspiracies and synchronic teleology

While constraints certainly allow a perspicuous statement of synchronically avoided or favored configurations, it is a mistake to argue that they are the exclusive prerogative of OT, and that therefore only it can formulate such statements. In fact, Vaux (2008: 29–30, fn 10) notes that by 1993 "most rule-based theories employed a suite of inviolable output constraints, such as the OCP, which were perfectly capable of generating conspiratorial effects".

One should also be skeptical of OT's driving assumption that cross-linguistically avoided configurations and cross-linguistically common processes all need to be manifest in some way in synchronic grammar, and that in general the theory of grammar is the appropriate locus of explanation for these typological generalizations. The perceived need to formulate all explanation in terms of synchronic teleology also fuels the OT account for conspiracies, and the critique of RBP based on them. We believe that in general the phenomena attributed to such teleology by OT are better accounted for within a diachronic/evolutionary approach to phonological typology, as pioneered by such works as Ohala (1972, 1975, 1981, 1989, 2005); Blevins (2004); and Ritt (2004) (see also section 7.3). On this view, the locus of typological explanation lies outside of the phonological grammar *per se*, and thus the argument from conspiracies against RBP is neutralized.

It is not obvious to us that the sort of synchronic teleology entertained in OT, even if empirically well-motivated, leads to an overall more deeply explanatory account of the nature of natural language phonologies than that available in a rule-based theory allied with the diachronic/evolutionary approach to typology. Notice that the teleological aspect of the OT approach to conspiracies has to come from the pre-existence of the markedness constraint implicated in a given conspiracy. As for why there should be such a markedness constraint in UG in the first place, the answer would presumably have to be "because that's what evolution did". But this is almost the same as the claim made by the alliance of RBP with diachronic/evolutionary phonology, with the following difference. In the OT case biological evolution would have to be appealed to – that is, the genetic basis for the faculty of language itself has evolved such that it contains the markedness constraint. In the RBP case, historical linguistic evolution would be appealed to, since by hypothesis there is no synchronic teleology encoded in UG itself: the appearance of a conspiracy, and all cross-linguistically avoided or favored configurations, emerges from the winnowing effect of human perceptual and articulatory mechanisms, coupled with the nature of the acquisition process. A form of natural selection is being appealed to in each case, and by no means is the OT approach obviously better in its coverage. Nor is it more likely to be true, for the reasons outlined above.

Along similar lines, we should not be too hasty to assume that a given set of processes is controlled by a synchronically active conspiratorial global constraint. As Kiparsky (1972: 190f.) states, the elements putatively implicated in conspiracies:

> are linguistically complex configurations, and rules eliminating or avoiding them are accordingly highly natural and occur frequently in the languages of the world. It is therefore only to be expected that there should be some languages in which several rules should eliminate or avoid these configurations, and that there should be languages in which no instances of these configurations appear on the surface. [. . .] What I am questioning, then, is whether there is any fundamental sort of difference between the cases in which just one or two rules reflect general phonological conditions of this type, and the cases in which several rules are involved, which would be termed a "conspiracy".

Kiparsky (1972: 191) adds that:

> concrete empirical differences are clearly also involved: Is there any evidence for a true "functional unity" of the rules in a conspiracy which would not simply be characterizable by their sharing a common target? Are there cases in which they are subject to parallel historical changes at some point in the development of a language? Are there cases in which apparently diverse changes in the rules of a language at some point in time can be shown to be consequences of the imposition of a single derivational constraint? Are there cases where the rules in a conspiracy have the same set of lexical exceptions? This would be strong evidence in favor of derivational constraints.

Kiparsky noted in 1972 that he had not found any cases of the sort he identified, leaving the reader to infer that the case for synchronic conspiracies had not been carried. To the best of our knowledge these questions have not subsequently been investigated; since 1993 it has simply been taken for granted that conspiracies exist, with no evidence for their existence in a synchronic grammar requested or provided.

A final concern with OT's attribution of conspiracies to the activity of synchronic markedness constraints is the fact that these constraints do not always appear to suffice to account for the conspiracies in question. The famous Yawelmani Yokuts conspiracy, for example, has been attributed by Heinz (2008) and others to the high ranking of a markedness constraint *Complex triggering avoidance of complex syllable margins, but *Complex on its own fails to account for Newman's observation that "no combination of the two syllabic types can result in a vowel cluster; all vowels must appear singly in Yokuts [; . . .] no syllable has an initial vowel" (1944: 27). Zoll (1993) proposes that the conspiracy actually aims to make all syllables be of the shape CV(X), which accounts for Newman's generalization but requires a conspiracy between *Complex, Onset, and *Superheavy in Zoll's analysis. Moreover, neither Heinz's nor Zoll's analysis accounts for all elements of the putative conspiracy, such as the fact that */CCC/ strings are banned in the underlying forms of morphemes (Kisseberth 1970) and "the initial syllable of [the underlying form of] bases is always open" (Newman 1944: 27).

One might object that the morpheme structure constraints proposed by Kisseberth and Newman are simply products of the history of the language, and need not be encoded in the synchronic grammar. This reasonable move leaves one wondering whether the entire Yokuts conspiracy might be solely historical. Blevins (2004: 47) suggests that claims about the synchronic activity of conspiracies can be tested by investigating the behavior of loanwords. To the best of our knowledge this has not been done by proponents of the Optimality Theoretic conspiracy analysis, but Gamble (1989) has found that Spanish loans in the closely related Wikchamni Yokuts do *not* employ consistent strategies to avoid word-initial CC-clusters (e.g. cruz 'cross' → kuluʃ but clavo 'nail' → la:wu), and Weigel (2005) has found that some Yawelmani loans from Spanish do allow complex onsets, such as escuela 'school' → eskwela.

When evaluating the phonological behavior of loanwords one must always be wary of the possibility that the words in question have come into the recipient language via one or more intermediate languages, as Gamble (1989) points out. At our current state of knowledge about Yokuts, though, it is fairly clear that a satisfactory case has not been made for a single markedness constraint synchronically producing conspiratorial effects.

7.2.1.1.3 Conclusions

In this light, it is premature to assert that OT offers a predictive advantage over RBP with respect to conspiracies. The synchronic existence of conspiracies has not been securely established, and in the Yokuts case the behavior of loanwords suggests that the internal and sociolinguistic histories of the language may play the key explanatory roles. The Yokuts case reveals moreover that even if a conspiracy *is* produced by the synchronic grammar, a single markedness constraint does not suffice in accounting for it; as in RBP, a conspiracy of players (rules in RBP, constraints in OT) is required to generate the desired patterns.

Even if it does turn out that conspiracies exist in non-trivial numbers, it is not clear that an account relying on the genetic substrate of the language faculty having evolved appropriate markedness constraints would provide the most insightful explanation of this. We find the Evolutionary Phonology position persuasive in this context: if independently required human perceptual, articulatory, and acquisitional mechanisms operating over time suffice to account for the appearance of conspiracies, then there is no reason to postulate additional synchronic machinery to duplicate the same effect.

7.2.2 Predictive differences suggested to favor RBP

Turning to predictive differences that have been suggested to favor RBP over OT, Anderson (2010) states that "some of the same issues that rule-based phonology dealt with (and at least largely resolved) have re-surfaced as serious challenges to the architecture of grammar generally assumed in constraint based theories". He singles out opacity as a particular challenge, and this phenomenon is the focus of our next section. We then briefly review "crazy rules", morpheme structure constraints, and the larger issue of locality vs. globality.

7.2.2.1 Opacity

The phenomenon of opacity is frequently cited as showing a need for a phonological architecture in which processes can interact in series (Kiparsky 2000; Clements 2000; McCarthy 2007; Vaux 2008; many others). Since serialism is one of the major theoretical fault-lines dividing RBP from Classic OT, the existence of opacity has often been taken as an important argument favoring RBP (indeed, RBP was the framework which allowed opacity to be discovered and investigated for the first time). The reason why has been rehearsed in many places in the literature; for this reason we provide only a brief illustration here. Following this illustration, we consider two issues which might seem to complicate the use of opacity as a test-bed for comparing RBP with OT, ultimately concluding that the evidence from opacity still comes down heavily in favor of RBP.

The notion of phonological opacity, and the surface diagnostics used to identify it, originate with Kiparsky (Kiparsky and Kiparsky 1971: 621–622; Kiparsky 1973: 79). The following definitions, taken from Kiparsky (1973: 79), are usually taken as canonical:

(11) A process P of the form A→B/C__D is opaque to the extent that there are surface representations of the form:
a. A in the environment C__D, or
b. B derived by P in environments other than C__D.

McCarthy (1999b) dubs the subcase in (a) *underapplication* opacity: a process fails to apply despite having apparently had an opportunity to do so. Subcase (b) is termed *overapplication* opacity by McCarthy, since it involves a process applying despite its usual conditioning environment not being met on the surface. Both of these classical subcases of opacity are predicted to exist by RBP, since both are easily expressible in terms of serial rule interaction.

A popular example of overapplication opacity occurs in Tiberian Hebrew, where a process of epenthesis crucially applies before a process that deletes glottal stops in syllable codas. In RBP terms, input–output mappings involving these two processes such as /daʃʔ/ → [dɛʃɛ] vs. /daʃʔ-o:/ → [daʃʔo:] can be derived with the three rules in (12) (Green 2004):

(12) Tiberian Hebrew counterbleeding in RBP

	UR	a. **/daʃʔ/**	b. **/daʃʔ-o:/**
Rule 1	Ø → [ɛ]/C_C#	daʃɛʔ	–
Rule 2	a → ɛ/__ C ɛ	dɛʃɛʔ	–
Rule 3	ʔ → Ø _ $]_\sigma$	dɛʃɛ	–
	SR	**[dɛʃɛ]**	**[daʃʔo:]**
		‘grass’	‘his/its grass’
		Gen1:11	(unattested)

In derivation (12a), Rule 1 counterbleeds Rule 3: if Rule 3 were to apply before Rule 1, it would bleed the latter’s opportunity to insert an epenthetic vowel.

In contrast, both types of opaque interaction pose problems for Classic OT analyses. In the case of underapplication opacity, easily analyzed as a case of counterfeeding rule ordering in RBP, there is no way to prevent process P from applying once its conditioning environment is met on the surface. Given the constraint ranking needed to cause the process to exist in the language in the first place, it will be forced to apply in any surface form where its structural description is met.[8]

Similarly, overapplication opacity of the sort dealt with by counterbleeding rule interactions in RBP poses insuperable difficulties for any account in terms of Classic OT. Consider how the counterbleeding interaction in (12) might be modeled in Classic OT (Green 2004):

(13) Tiberian Hebrew counterbleeding in Classic OT

/daʃʔ/	*CC]Σ	*ʔ]Σ	Dep-IO(V)	Max-IO(C)
[daʃʔ]	*!	*		
[dɛʃɛʔ]		*!	*	
☞ [dɛʃɛ]			*!	*
☜ [daʃ]				*

The key in (13) is the harmonic bounding of the desired counterbleeding candidate [dɛʃɛ] by the transparent bleeding candidate, *[daʃ]: because [dɛʃɛ] violates all of the same constraints as *[daʃ], and also violates additional constraints, there is no ranking of the constraints that can select the counterbleeding candidate over its transparent competitor.

The reason stems from the very logic of constraint interaction in the theory. Processes can only exist when a given (set of) markedness constraint(s) outranks a certain faithfulness constraint. Therefore, processes are predicted not to apply if the markedness violation which would trigger them is not present on the surface – there is no room in Classic OT for gratuitous violation of a faithfulness constraint. The issue created by overapplication opacity is that, by its very definition, it inevitably involves such a gratuitous faithfulness violation.

Since both of these types of opacity are a productive part of the phonologies of many languages, there can be no doubt that they constitute a strong argument in favor of RBP against Classic OT. However, it should be pointed out that this argument from opacity turns on the parallelist nature of Classic OT in combination with its commitment to ranked constraints, and not on its constraint-based nature alone. Various ways of allowing OT evaluations to interact serially are currently being actively pursued. One of these allows an OT grammar to evaluate entire candidate derivations (Optimality Theory with Candidate Chains, q.v. McCarthy 2007; Wolf 2008). Another combines an Optimality Theoretic phonology with the stratal morphophonological architecture of Lexical Phonology and Morphology, allowing the output of one stratum to serve as input to a subsequent one in a serial fashion (LPM-OT; see Kiparsky 2000; Bermúdez-Otero 2013, forthcoming, this volume).

These models have been independently criticized (see Kiparsky 2011 on OT-CC; Embick 2010: Ch 1, 6, and 7 on LPM-OT), and are subject to many of the other criticisms of OT raised elsewhere in this chapter. Nevertheless, if a serial variant of OT proves viable in the long run, the relevance of opacity phenomena as a way of contrasting the predictions of RBP with those of OT may be put into doubt.

A separate issue which seems to complicate the use of opacity as a test-bed for RBP and Classic OT has recently been raised by Baković (2007, 2010). Baković argues that traditional definitions of opacity, properly applied, include certain phenomena which do not receive a satisfying analysis under RBP, unless it is supplemented by additional mechanisms. The importance of Baković's argument is that, if indeed RBP does not have a fully unified analysis of opaque phenomena, it can no longer be taken to be inherently superior to Classic OT as an approach to opacity. Before discussing the details of the problematic phenomena identified by Baković, we would point out that the structure of this argument contains an important flaw: Classic OT and most of its variants lack *any* satisfactory mechanism for generating or learning (large classes of) counterfeeding and counterbleeding systems. It follows that RBP is to be preferred even if Baković is correct that RBP cannot achieve a unified analysis of all opaque phenomena – having a non-uniform analysis is better than having no analysis.

The two cases of opacity which Baković points out as problematic for RBP are blocking and cross-derivational feeding. We discuss these in turn.

Blocking includes a number of effects which are well-known in the phonological literature, but had not previously been widely acknowledged to fulfill the definition of opacity, including: blocking of one process by another via the Elsewhere condition; non-derived environment blocking (where processes fail to apply outside of derived environments); and do-something-except-when blocking (where processes fail to apply exactly where they would create a configuration which is banned on the surface). As Baković astutely observes, all of these types of blocking fall under the heading of underapplication opacity given Kiparsky's canonical definition. Baković points out that all of these phenomena require ancillary conditions beyond pure rule interaction in RBP in order to account for them. While correct, this argument is subject to our earlier objection: this does nothing to alleviate Classic OT's inability to deal with counterbleeding and counterfeeding-on-environment.

Cross-derivational feeding is Baković's term for a class of interactions which he argues cannot be captured in RBP without loss of generalization. A simple example comes from the interaction of epenthesis with Voicing Assimilation in English phonology. As is well known, English inflectional suffixes with underlying forms /z/ and /d/ are devoiced to [s] and [t] when they follow a voiceless obstruent. However, this process is bled by a process of epenthesis which applies to separate a sequence of two (near-)identical consonants. This interaction is illustrated for the English regular plural inflection in (14):

(14) Epenthesis and Voicing Assimilation in English

a. Rules

Epenthesis (Ep)	Ø → ə/[−son, αPlace] _ [−son, αPlace]
Voicing Assimilation (VA)	[−son] → [−voice] /[−voice] _

b. Derivations

	'days'	'dads'	'plates'	'cakes'	'dishes'	'buses'
UR	/dɛɪ-z/	/dæd-z/	/plɛɪt-z/	/kɛɪk-z/	/dɪʃ-z/	/bʌs-z/
Ep	–	–	–	–	dɪʃəz	bʌsəz
VA	–	–	plɛɪts	kɛɪks	–	–
SR	[dɛɪz]	[dædz]	[pʰlɛɪts]	[kʰɛɪks]	[dɪʃəz]	[bʌsəz]

The core of Baković's observation is as follows (the ensuing discussion is based on Fruehwald and Gorman 2011: 37). Epenthesis applies in this instance to (in Baković's opinion) break up sequences of sounds which are "too similar". Baković claims that the only feature that doesn't "count" in determining whether two sounds are different or not is [voice] – if two sounds differ in [voice] but are otherwise the same, epenthesis still applies. Crucially, [voice] is exactly the feature manipulated by the assimilation rule. Baković argues that this dual role of the [voice] feature is not a coincidence, and is a generalization which should be captured by the grammar. This relationship can be captured in OT, thanks to the global nature of its computations. This is because the constraint which favors epenthesis (the anti-gemination constraint NoGem) interacts with the constraint favoring agreement of voicing in word-final obstruent clusters (Agree(Voice)). Provided that these two markedness constraints are undominated[9] and are not ranked with respect to each other, the correct result is captured:

(15) Cross-derivational feeding in OT

/bʌs-z/	NoGem	Agr(Voi)	Dep(V)	Id(Voi)
a. bʌsz		*!		
b. bʌs:	*!			*
☞ c. bʌsəz			*	
d. bʌsəs			*	*!

This interaction is termed "cross-derivational feeding" because the ultimate reason for the application of epenthesis is that there *would be* a geminate if Agree(Voice) alone were satisfied (see candidate (15b)) – that is, the winning candidate can only be determined by reference to the output of a counterfactual derivation in which assimilation does apply. RBP grammars disallow such cross-derivational comparison, and so can offer no grammatical account of this interaction. Furthermore, Baković points out a strong prediction made by the OT approach to cross-derivational feeding which is not made by the RBP approach: note that candidate (15d), in which both epenthesis and Voicing Assimilation have applied, is harmonically bounded by candidate (15c).[10] Thus, OT predicts that it should be impossible for assimilation to counterbleed epenthesis cross-linguistically, a prediction which Baković (2007) defends against the only known apparent counter-example (from New Julfa Armenian, see Vaux 1998; see Fruehwald and Gorman 2011: 44 for a telling reply to Baković's defense).

While interesting, there are at least four important problems with Baković's argument from cross-derivational feeding. The first is that the OT approach in (15) rests its most impressive typological prediction on OT's general inability to cope with counterbleeding. This is, to say the least, a dubious move: it makes a small virtue out of an otherwise crippling liability for the theory. The next two problems are pointed out by Fruehwald and Gorman (2011). They note that epenthesis does not always systematically ignore all and only the features relevant to assimilation, as predicted by Baković's approach. In fact, there are exceptions even in English, with respect to the alveolo-palatal affricates: note that applying Voicing Assimilation to /dɪʃ-z/ does not yield a geminate. They further note that the interaction of Voicing Assimilation and epenthesis in modern English can be given a plausible extragrammatical explanation in terms of its historical origin as a syncope rule which was blocked where it would have given rise to a geminate (a blocking which Fruehwald and Gorman 2011: 38–41 attribute to homophony-avoidance pressures operating over time). There is thus no need to capture the generalization at the level of the grammar, and the apparent disadvantage of RBP evaporates.

The fourth problem is that Baković's strong prediction that epenthesis cannot counterbleed assimilation is actually false – such grammars do in fact appear to exist: for example, epenthesis has been argued to counterbleed laryngeal assimilation in some English idiolects

(Anderson 1973), the Armenian dialects of Maragha and Nor Nakhichevan (Vaux 2016), Japanese (Davis and Tsujimura 1991), and some northern Greek dialects (Newton 1972: 207).

Overall, then, it seems that the phenomenon of opacity still favors the predictions of RBP over those of at least Classic OT, and that Baković's arguments to the contrary are not convincing.

7.2.2.2 Other predictive differences claimed to favor RBP

Anderson (2010: 610) states that cases:

> seem to exist in which the specific changes through which a language achieves conformity with a general constraint on surface forms do not follow directly from the content of the constraint (together with other interacting generalizations). In such a case, something like a re-writing rule might be necessary, as a supplement to the constraint system – a notion which is clearly antithetical to the basic philosophy of OT.

Here he appears to be referring to "unnatural" or "crazy" processes, that is (to invert Donegan and Stampe's (1979: 126) characterization of Natural Phonology), ones which are not a natural reflection of the needs, capacities, and world of its users (Bach and Harms 1972; Anderson 1981).

One such example is consonant epenthesis, which in many languages inserts a segment that appears to result from historical reanalysis of an earlier consonant deletion process, rather than from minimal departure from the underlying representation in order to satisfy markedness constraints, as we would expect in OT. Cases like this abound in the world's languages, most famously with English *r* but also with Mongolian *g*, Dominican Spanish *s*, and many others (Vaux and Samuels 2017).

The problem for all varieties of OT in this context arises in our opinion from the fact that OT descends philosophically from Natural Phonology (NP), one of whose tenets is that phonology is essentially natural in the sense defined earlier. In NP this naturalness is captured for the most part in the idea that UG provides a set of "processes" such as Final Devoicing; their initial state in the acquisition process is to be active, but the learner can suppress them upon sufficient exposure to primary linguistic data (PLD) showing that the process in question is not active in that language. In the case of Final Devoicing, for example, exposure of the learner to surface forms showing a voicing contrast in word-final obstruents would suffice to deactivate the process.

In OT the analog is a universal set of constraints with a particular initial ranking (typically some form of $M \gg F$) thought to reflect/produce the same phenomena encoded as processes in NP. Parallel to the deactivation of natural processes in NP, markedness constraints in OT can be demoted below relevant conflicting faithfulness constraints when the learner observes output forms that violate these markedness constraints.

NP and OT diverge, though, when it comes to *un*natural phenomena. While NP asserts that phonology is fundamentally natural, it acknowledges the existence of another module of operations that precede Processes in the derivation. These operations, called Rules (analogous to lexical rules in Lexical Phonology and morphophonemic rules in traditional phonology), are not provided by UG but instead have to be constructed by the learner. It is this class of operations, and particularly the subset classified as unnatural by Anderson (1981), that has been abandoned by OT, the claim being that one cannot for example have a phonological process that inserts a random consonant – if a language has a consonant insertion process, the choice of consonant must be the one that minimally deviates from the underlying form while optimally satisfying the constraint ranking of the language.

Technically the machinery of OT *is* capable of generating insertion of any given consonant, given that no relevant bounds have been placed on the inventory of constraints in UG. There is nothing to stop us from postulating additional markedness and faithfulness constraints that will suffice to produce [g]-insertion between long vowels in Mongolian, for example. On a philosophical level, though, most practitioners of OT are not willing to allow for this degree of generative power,[11] which leaves us with a predictive difference between RBP (which allows for unnatural operations such as Mongolian [g]-insertion) and OT (which philosophically at least does not).

Another area where the computational power and the philosophical predilections of OT differ involves morpheme structure constraints, or more generally constraints on the form of underlying representations. Optimality Theoretic constraints have the ability to inspect underlying representations (this is an essential component of faithfulness constraints, for example), and there is currently no substantive proposal for the set of constraints contained in UG, so there is no formal or computational reason why markedness constraints legislating the form of input representations should not exist. As with unnatural processes, though, the philosophical inclinations of most supporters of OT do not exactly align with what the formalism allows. In this case, adherents of OT generally believe that UG does not contain constraints on underlying forms, a principle called Richness of the Base; the reasoning generally cited is that OT allows one to capture all of the significant generalizations about a language without recourse to such constraints, and therefore by Occam's Razor they should not be postulated.

Nevins and Vaux (2007) argue, though, that there is extensive cross-linguistic evidence for precisely this sort of constraint on underlying representations. Becker and Gouskova (2016) give experimental evidence that speakers extend generalizations over underlying forms in Wug tests. Rasin and Katzir (2014) demonstrate, moreover, that under standard principles of acquisition in OT, language-specific constraints on underlying representations are required in order to avoid treating some systematic gaps as accidental and some accidental gaps as systematic.

A final class of predictive differences where the facts appear to support RBP over OT involves locality effects that OT's globality of optimization makes difficult to capture. A range of such cases have already been discussed by Paster (2006) and Embick (2010); we therefore limit ourselves to a somewhat different case involving localized iterative clash deletion.

Meeussen's Rule (MR) avoids strings of three adjacent High tones; when handed an /H-H-H/ string by the morphology, languages with MR produce either [HLL] (as in Ganda; Hyman 1989) or [HLH] (as in Shona; Odden 1980). We provide a Ganda example in (16).

(16) Ganda (Hyman 1982 and Hyman 2001 fig.14)

```
H     H     H
|     |     |
ba -  li -  lab - a  'they will see'
|     |     |
H     L     L
```

The two variants of MR are traditionally analyzed as OCP-driven High-tone delinking, applying iteratively either from left to right (as in Ganda) or right to left (as in Shona). In an OT framework, though, we only expect to find the HLH treatment and not the HLL treatment, as the latter involves an unnecessary second Max violation (Odden 2008: 71). We thus have a predictive difference between RBP, in which configurations are evaluated and operated on locally, and OT, in which the same is done globally: the former predicts the existence of both types of MR, where the latter generates only the type of MR found in Shona.

A parallel to the behavior of MR in Bantu surfaces in the Abkhaz stress system (see also Vaux 2008: 39–40, on which the ensuing discussion is partly based). The core stress system in Abkhaz is governed by Dybo's Rule, which can be formulated as in (17):

(17) Dybo's Rule (Dybo 1973, 1977, 1978; cf. Hewitt 1979, 1989; Spruit 1986: 38; Trigo 1992; Kathman 1992, 1994, 1996)
Assign word stress to (i) the leftmost underlyingly accented element (ii) not followed by another accented element; otherwise (iii) stress falls on the final element.[12]

We illustrate the workings of Dybo's Rule with the forms in (18)–(19). In the notation employed here, underscored "x" represents a lexically accented element and an acute accent "x́" represents an element that is stressed in the surface form.

(18) Nominal root: /madza/ 'secret' (Spruit 1986: 42)
a. [á-madza] 'DEF-secret'
b. [madzá-k'] 'secret-INDEF'

(19) Verbal roots (Spruit 1986: 46)

	Unaccented root			**Accented root**	
a.	á-pʰa-ɾa	'jump'	d.	a-pʰa-ɾá	'pleat'
b.	á-fa-ɾa	'eat'	e.	a-ja-ɾá	'lie down'
c.	á-tʰa-ɾa	'give'	f.	a-tsʰa-ɾá	'go'

We can see the workings of condition (i) of Dybo's Rule in (18a) and (19a–c). In (18a), the lexically accented definite prefix /a-/ is followed by the unaccented root /madza/; by dint of (17i) the /a-/, being the leftmost underlyingly accented element, receives the surface stress. The forms in (19) are verbal infinitives, which are constructed from the root by prefixing the accented definite morpheme /a-/ and suffixing the accented infinitive ending /-ɾa/. In the forms in (19a–c) the verbal roots are underlyingly unaccented; the lexical accents of the prefix and suffix are therefore non-adjacent, and condition (i) of Dybo's Rule therefore holds again, assigning stress to the leftmost of these accents.

Using Halle and Idsardi's (1995) formalism, we can derive the basic Abkhaz system in RBP via the operations in (20).

(20) Abkhaz stress in the Halle–Idsardi formalism
i. Project stress-bearing elements
ii. Project a right bracket) for all lexical accents
iii. Line 0 Edge Marking: LLL
iv. Clash Deletion:) → Ø/_ *) [iterative, L→R]

v. Project rightmost element of Line 0 feet to Line 1
vi. Project leftmost element of Line 1 feet to Line 2

The conflicting directionality identified by Dybo results from Left vs. Right headedness on Lines 0 and 1 respectively (20v, 20vi), and the iterativity and directionality via (20iv) (cf. Howard 1972).

In OT, on the other hand, it is difficult to derive the equivalent of iterative left-to-right clash deletion, for the same reasons we saw with Meeussen's Rule earlier. The Classic OT tenets of globalism/parallelism and minimal violation favor outputs which do the global minimum necessary to avoid stress clash, which harmonically bound the desired winners with their greater number of clash deletions. OT therefore predicts that Abkhaz should delete as few underlying lexical accents as possible, yielding a Shona-style system (which would produce forms like /a-pʰa-ɾá/ 'pleat' → *[ápʰaɾa]) and not the attested Ganda-type system (which produces forms like /a-pʰa-ɾá/ → *[apʰaɾá]).

Versions of OT that allow staged computation sensitive to morphological structure (e.g. Stratal OT or in a sense Orgun's (1996) Cyclic OT) can deal with at least a subset of the cases where no more than one deletion happens per morpheme, but Stratal OT fails with forms involving more than one deletion per stratum and Cyclic OT fails with forms involving more than one deletion per morpheme. Harmonic Serialism has the power to generate the desired outputs, but cannot rule out equally harmonic outputs produced by derivations that do not apply clash deletion in L→R order.

What we conclude from the behavior of Meeussen's Rule in Ganda and Dybo's Rule in Abkhaz is that the phonological component of the human language faculty requires the ability to execute operations in a non-optional, (process-specific) directional, local manner. Theories designed to be unable to carry out such computations and/or select the outputs of such computations as the exclusive winners under Eval face a serious challenge accounting for the relevant empirical phenomena.

We have seen in this section a representative selection of predictive differences that have been suggested in the literature to favor RBP over OT. In the case of opacity, RBP straightforwardly generates all attested types of counterfeeding and counterbleeding, whereas all mainstream forms of OT other than OT-CC struggle to account for the complete range of opacity effects (see McCarthy 2007 for an insightful review of the relevant theories and phenomena). Attempts by Baković to label additional phenomena as opacity effects are ultimately unsatisfying, as they brush under the rug the fact that OT cannot generate counterbleeding and counterfeeding-on-environment effects, and downplay the fact that RBP does in fact possess mechanisms to generate the effects in question.

We have seen that unnatural processes can be generated in some cases by the machinery of OT, but the philosophy of the theory precludes the existence of such processes, which then runs afoul of the fact that they are abundantly attested in the languages of the world. The same line of reasoning holds for morpheme structure constraints: nothing in the architecture of OT prevents them from existing, but proponents of the model have nonetheless made Richness of the Base a central concept in practice, which poses significant problems both empirically and learning-theoretically.

Finally, iterative directional clash deletion processes of the sort we see in Ganda and Abkhaz are wrongly predicted not to exist by most varieties of OT. OT-CC is able to generate forms of the sort we observe in these languages, but these are predicted to tie with globally optimal forms, contrary to fact.

7.3 An outstanding RBP-internal issue: the content of Universal Grammar

This section highlights an open issue internal to the theory of RBP: the content of UG. Another major current issue within RBP is morphology–phonology interaction and the architecture of the grammar; see Newell (this volume) for an overview of current work on this topic.

Just under three decades ago, McCarthy (1988: 84) was able to write that "[t]he goal of phonology is the construction of a theory in which cross-linguistically common and well-established processes emerge from very simple combinations of the descriptive parameters of the model". Today, the view that generalizations about typological frequency should be made to fall out from the primitives made available by UG has been challenged from a number of perspectives.

One line of criticism has been conceptual in nature: the aim of linguistic theory is to characterize the human language faculty. The core question of interest is therefore what is a *possible* language, not what is a *probable* language (see Hale and Reiss 2008; Newmeyer 2004). Moreover, of the unattested languages, only a subset of them will be unattested because they are not computable by the language faculty – others may be computable in principle, but unattestable for independent reasons. To take an extreme example, Hale and Reiss (2008: 192) point out that there are no linguistic representations in human languages that contain physical bananas, but this is clearly for extragrammatical reasons and presumably requires no UG-based explanation.

Another line of criticism stems from the emergence of new kinds of explanation for why certain processes are more "cross-linguistically common and well-established" than others. Every currently existing phonological grammar is the result of a baby acquiring its ambient language on the basis of the linguistic output of its care-givers and peers. Since this means that grammars are not transmitted directly, but rather "reassembled" by each individual acquirer, there is a good chance that transmission from one generation to the next will be imperfect, leading to language change. Work on phonetics has shown that, for fundamentally articulatory and perceptual reasons, certain such language changes are more likely to happen than others (Blevins 2004; Ohala 1971, 1972, 1975, 1981, 1983, 1997, 2005).[13]

These developments have raised the question of whether the notion of *markedness* has any status in UG. This question is intimately related to a second one – namely, how sensitive is the phonological component to the fact that its features have phonetic content? If the grammar itself disfavors phonetically marked structures (i.e. ones which are difficult to articulate and/or perceive), then it clearly must be sensitive to the phonetic substance of distinctive features. On the other hand, if the effects of phonetic markedness are purely extragrammatical, then it could be that phonology itself is "substance-free". It is important to point out that these issues arise in any theory, and in fact this discussion is independent of the rule-based nature of the framework. Within RBP, SPE (especially chapter 9), NP (Stampe 1973, 1979), and much work in Autosegmental Phonology can be regarded as substance-ful approaches to phonology; see Hale and Reiss (2008) and Samuels (2009) for extended defenses of substance-free versions of RBP. Similarly, although most Optimality Theoretic literature from Prince and Smolensky (1993) onwards is in the substance-ful camp, Blaho (2008) proposes a substance-free implementation of OT.

One empirical front in this debate focuses on experimental phonetics and the predictions of a sound change-based theory for phonological typology. Do the set of phonetically natural mishearings and mispronunciations coincide with the set of common sound changes, and does the set of attested sound changes correctly predict the set of attested/unattested synchronic phenomena? Blevins (2004) argues that the answer to both of these questions is "yes". Kiparsky (2006); de Lacy (2002, 2006a, 2006b); and de Lacy and Kingston (2006) argue that the answer to the second question is "no", on the grounds that certain phonetically natural and common sound changes, if chained together, are predicted to yield certain synchronic phenomena which turn out not to be attested. Miller et al. (2016) reply that while the changes cited by Kiparsky, de Lacy, and Kingston may be common individually, this does not mean that the chains of them needed to produce the anomalous patterns are expected to be common. Furthermore, Miller et al. show that many of the chains of changes suggested by Kiparsky actually yield commonly attested patterns, rather than the unattested ones claimed. This is a rich empirical domain which will continue to play an important role in the debate on how "substance-ful" phonology is (see also Reiss, this volume).

Another important domain in this respect has been language acquisition. If phonology is substance-free, then formal complexity and robustness of data should be the only determinants of how "easy" to acquire a given pattern is. All else held equal, a phonetically unnatural pattern should be no more difficult to acquire than a natural one (as we will see in section 7.4, however, all else might well not be equal). On the other hand, if phonology is sensitive to phonetic substance and rules out certain unnatural patterns (for instance, via harmonic bounding in an OT framework), then phonetically unnatural patterns should be impossible to learn, whereas natural patterns which are otherwise equal in terms of formal complexity and robustness in the PLD should be easy to acquire.

Recent years have seen an explosion of experimental work on this issue, including Moreton 2008; Finley 2008; Finley and Badecker 2009; Pater and Tessier 2003; Wilson 2003; Peperkamp et al. 2006; Berent et al. 2007; Berent et al. 2008; Berent et al. 2009; Hayes et al. 2009; Pycha et al. 2003. See Hayes and White 2013 for an overview).

A particularly common paradigm involves the learning of artificial languages, whether by infants (Saffran and Thiessen 2003; many others) or by adults (Wilson 2006; many others). Another methodology involves comparing observational studies of the course of acquisition of phonology in L1, seeking correlations between phonetic markedness/typological rarity and lateness of acquisition (Buckley 2002).

The picture that emerges from this literature is not easy to interpret, however. As Hayes and White (2013: 47) put it: "Our reading of this literature is that the evidence is quite mixed and gives no comfort to advocates of either of the two possible extreme positions (all constraints are a priori knowledge/all learning is purely inductive)". Hayes and White (2013: 47–48) survey a number of artificial language-learning studies which show that unnatural rules can indeed be learned – in some cases more slowly than natural ones (as found by Schane et al. 1974/1975), and in some cases with no apparent difference (for example, Pycha et al. 2003, who found that adults acquired the typologically rare pattern of consonant harmony just as easily as the far more common process of vowel harmony). In addition, Buckley (2002) found that studies on the order of L1 acquisition of different processes cross-linguistically show no particular evidence that "natural" processes are easier to learn. On the other hand, they are also able to cite a range of experiments which support at least some role for phonetic naturalness. For instance, Wilson (2006) found that adult learners extended a palatalization process of ke → tʃe to ki → tʃi, despite only having independent evidence in their training data for ke → tʃe. However, participants trained on data which only carried evidence of a ki → tʃi process did not extend this rule to ke → tʃe when tested. This asymmetry accords with what phonetic naturalness would predict: high vowels are more likely to induce palatalization than mid vowels. The same asymmetry is reflected in the typological record – if /e/ triggers palatalization in a given language, then so will /i/, but the converse does not hold (Hayes and White 2013: 48).

Summing up this mixed picture, Hayes and White (2013: 49) note that "[in] several of the experiments just cited, the findings support a *bias* effect: the unnatural patterns are learnable but take longer to learn, or yield weaker experimental effects than comparable natural patterns". This conclusion clearly rules out certain versions of the substance-ful approach (for example, it is incompatible with the sort of absolute absence of unnatural patterns which would be predicted by a harmonic bounding-based OT explanation), while being compatible with other substance-ful approaches. However, it does not rule out substance-free approaches either. The reason for this is that a learning bias need not emerge from the UG primitives from which phonological grammars are assembled. Instead, the learning bias could be part of the evaluation metric in the sense of Chomsky (1955, 1957, 1965) – i.e. it is a way in which the Language Acquisition Device chooses amongst candidate grammars which account for the PLD, not part of those grammars themselves. The intuition here could be informally stated as follows.

(21) Naturalness Bias (presumably one of many in the evaluation metric)
Given a choice between two or more hypothesized rules which account for the PLD, choose the one which yields phonetically natural outputs.

Given something like (21), the experimental results can be explained even if the grammar itself is substance-free.

This section has reviewed research on an open issue which must be faced in all phonological frameworks, including RBP: that of the content of the phonological component of UG. We have focused on the question of whether UG itself contains any direct encoding of the notion of "markedness". While most mainstream formal phonology through to the early 1990s was committed to a positive answer to this question, we have seen that the last two decades have given rise to an alternative position, according to which the generalizations covered by "markedness" emerge from the effects of articulatory, perceptual, and learning biases operating over time through language change. This latter position, if correct, makes it plausible that phonology itself is "substance-free" (Hale and Reiss 2008). We have attempted to outline the major empirical and conceptual parameters of this debate, which is ongoing.

7.4 Future prospects

In this section we outline a few areas where we believe that phonological research could make significant headway, regardless of one's theoretical predilections.

- **Language in other modalities**
  Though many theoretical linguists have believed for some time now that the human language faculty (including its phonological component) is modality-independent, phonological research still focuses almost exclusively on the medium of oral speech and acoustic processing thereof. As discussed in the previous section, this opens the door for confusions of speech-specific effects with general modality-independent phonological structures. In OT, for example, many of the constraints proposed to form part of Universal Grammar are specific to the oral/aural modality, such as *g "[g] is prohibited" (McCarthy and Prince 1995: 105), *a "assign one violation mark for each [a] in the

output" (Rosenthall 2006), or IDENTF1 "the first formant values of the input and the output must be identical" (Flemming 2002: 34).

We know from the study of sign languages, though, that language in non-speech modalities can also show phonological patterning and processes (Stokoe 1960; Brentari 1990, etc.), suggesting that phonology, like syntax, is modality-independent (and hence, in our opinion, part of the grammar/computational system that arguably forms the core of the human language faculty). As in our opinion there are at least hundreds and likely thousands of sign languages yet to be studied, this represents a rich test-bed for the investigation of what is universal and what is not in the phonological component of the language faculty.

- **Language acquisition**
Though the importance of conceptual issues and empirical data in language acquisition is already widely recognized by phonologists, the types of questions investigated tend to be relatively superficial: In what order are segments and syllable types acquired? How are phonotactics learned? More promising in our opinion is the use of (both normal and abnormal) acquisition data to investigate more abstract workings of phonological and general computation: What kinds of phonological hypotheses are entertained by learners? What kinds of phonological processes do and don't transfer from one's native system when learning a second language? Can learners acquire opaque phonological generalizations, and if so, how? The answers to these and similar questions have significant implications for competing phonological theories. For example, Ferguson (1992: 487–488, citing Dunn 1983; Elbert et al. 1984; Gierut 1985; LSHSS 1988; Tyler et al. 1987) points out that generalization in the course of first- and second-language phonological development ("when intervention by the therapist or teacher focuses on a particular pronunciation problem of the learner, improvement in the trained pronunciation takes place, and then related improvements are observed in untrained features of pronunciation") is a "valuable and under-utilized tool of phonological research [because] the range of atypical, idiosyncratic phonological systems and the possibilities of phonological mismatches across languages are so great". RBP encodes the cognitive phenomenon of generalization directly, in the form of rules. OT on the other hand normally employs demotion- and (in some variants) promotion-based learning algorithms, which do not encode rule-like generalizations directly and are typically taken to change the grammar gradually. These RBP and OT views of learning make strikingly different predictions about what sorts of intermediate systems should surface in the course of acquisition, and what sorts of systems are and are not learnable.

The acquisition of opaque phonological interactions has been investigated to a certain extent, both empirically (Jesney 2011; Dinnsen and Gierut 2008) and theoretically (Tihonova 2009; Tessier 2015). These studies typically operate within an OT perspective, without comparison to RBP and ignoring problems that the phenomena adduced pose for existing OT learning algorithms (Tessier 2010), a notable exception being the work of Andrew Nevins (e.g. Nevins and Braun 2009; Nevins 2010). Limiting one's purview to OT (or to RBP, for that matter) necessarily restricts the questions one asks, and the interpretation of one's findings. It is our opinion that investigation of acquisition with an eye towards identifying and testing predictive differences between theories offers more hope for progress.

One development in the field that may facilitate rapprochement between RBP and OT, and between rationalist and reductionist perspectives more generally, is the rise of sophisticated mathematical techniques for modeling the acquisition process, notably involving information theory and Bayesian probability (see e.g. Hayes and Wilson 2008; Goldsmith and Riggle 2012). The first two of these had significant effects on Jakobson and Halle and hence on seminal phonological works such as Jakobson et al. (1952) and Chomsky and Halle (1968), but their influence on phonological theorizing attenuated rapidly until these lines of thinking were (re-)introduced from computer science and statistics in recent years. Both Bayesian and information-theoretic models offer the prospect of integrating leading ideas from Chomskyan research (e.g. Bayesian priors are comparable to analytic biases in UG that constrain the hypothesis space) and more empirically and probabilistically oriented models (e.g. the maintenance and weighting of competing hypotheses in Bayesian analysis is in principle compatible with variationist theories such as Yang 2002).

7.5 Conclusions

We have tried to emphasize in this chapter the importance of focusing on predictive differences between theories (particularly, though not exclusively, RBP and OT), as this method reduces the likelihood of falling into local optimality traps (ones where the linguist reaches the best possible solution given a particular set of theoretical assumptions, but misses a superior analysis made possible in a competing framework). We have suggested that, contrary to what some have claimed in the literature, leading implementations of RBP and OT do in fact reveal predictive differences. While the weight of these supports RBP overall, we hold out hope that the best insights of both perspectives may eventually be integrated in a more robust model of the phonological component.

Notes

1 Cf. McCarthy (2002: 149): "the central thesis of OT is that a grammar is a language-particular ranking of violable, universal faithfulness and markedness constraints".

2 The generalized form of OT evaluated here, following Vaux (2008: 30–31, fn 11), represents our best attempt:

> to capture what [is] essential to the [theory], eliminating the inconsistencies and the debilitating unclarities of the various approaches that are developed in the literature. As an interpretation, it might be incorrect; but to reject attempts at such interpretation is pointless, since the only alternative is to reject what exists as inconsistent and vague, overlooking the important insights embedded in it.
>
> *(Chomsky 1967: 110)*

3 Lin therefore also punishes metathesis.

4 However, he conjectures that "a fixed ranking of ObsPlace Ident≫ NasPlace Ident" (p. 14) may be behind the generalization. See section 7.3 for a general critique of such importations of typological markedness hierarchies into synchronic constraint rankings, which are a commonplace of the OT literature (e.g. Prince and Smolensky 1993 on sonority constraints), and for an alternative approach.

5 Pater assumes that [nasal] is a unary feature.

6 Such root/affix asymmetries often err towards preserving faithfulness in lexical morphemes. For this reason separate faithfulness constraints have been postulated (by Pater and also by McCarthy and Prince 1995; Urbanczyk 1996) to deal specifically with root-internal faith violations.

7 We have corrected some errors in Pater's Quechua glosses, which have no bearing on Pater's discussion.

8 Strictly speaking, this is only true in cases where the counterfeeding effect is produced by changing the environment surrounding the would-be-affected segment, as in the present example. Another

type of counterfeeding involves processes applying to the same segment, such that one should feed the other, but fails to. Such *counterfeeding-on-focus* has been suggested by McCarthy (1999) to be amenable to an OT treatment employing faithfulness constraints which punish "too great" a change. Counterfeeding-on-focus therefore cannot be used to contrast the two theories, and for this reason we ignore it here.

9 A problem for this analysis is that NoGem cannot in fact be undominated in English, which allows Level II and Phrase Level geminates (e.g. un-known, pen name, one nation).

10 We are not convinced that this is a case of harmonic bounding. Consider the English voicing example, which is isomorphic to Baković's Lithuanian example, but with different constraints. Candidate (15c) does not in fact harmonically bound (15d), once one considers a more complete set of the constraints in the universal set Con. Many of the constraints in Con are actually violated by (15c) and *not* (15d), such as whatever constraint(s) make [z] more marked than [s] (e.g. *Lar in Lombardi's (2002) theory), and whatever constraint triggers final devoicing (e.g. *VcdObs$_{]PWd}$ in Staroverov's (2010) theory). Though Baković's formulation of the situation in terms of harmonic bounding appears to be unwarranted, his conclusion that Classic OT is not capable of generating such systems is nonetheless correct.

11 Advocates of substance-free OT such as Blaho (2008) and Iosad (2012) would be possible exceptions to this generalization.

12 Spruit (1986) discusses several classes of exception to generalization (iii) that we will not consider here as they are not germane to the point of this chapter.

13 Although this line of work is indeed relatively new in contemporary linguistics, Blevins (2004: 79) notes that there is an antecedent for both Ohala's and her own research program in the work of Baudouin de Courtenay (1910/1972: 267), who stressed "the importance of errors of hearing [. . .] as a factor of change".

References

Altenberg, Evelyn P. and Robert Vago. 1983. Theoretical implications of an error analysis of second language phonology production. *Language Learning* 33: 427–447.

Anderson, Stephen. 1973. Remarks on the phonology of English inflection. *Language & Literature* 1.4: 33–52.

Anderson, Stephen. 1981. Why phonology isn't "natural". *Linguistic Inquiry* 12.4: 493–539.

Anderson, Stephen. 2010. Phonology. In Patrick Hogan, ed., *The Cambridge Encyclopedia of the Linguistic Sciences*, pp. 609–612. Cambridge: Cambridge University Press.

Bach, Emmon and Robert T. Harms. 1972. How do languages get crazy rules? In Robert P. Stockwell and R.K.S. Macauley, eds., *Linguistic Change and Generative Theory*, pp. 1–21. Bloomington: Indiana University Press.

Backley, Phillip. 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Baković, Eric. 2007. A revised typology of opaque generalizations. *Phonology* 24.2: 217–259.

Baković, Eric. 2010. Opacity and ordering. In Goldsmith, John, Jason Riggle and Alan Yu, eds., *Handbook of Phonological Theory*, second edition, pp. 40–67. Malden, MA: Wiley-Blackwell.

Baudouin de Courtenay, Jan. 1972. *A Baudouin de Courtenay Anthology: The Beginnings of Structural Linguistics*. Edited by E. Stankiewicz. Bloomington: Indiana University Press.

Becker, Michael and Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47.3: 391–425.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel A. Moreno and Paul Smolensky. 2008. Language universals in human brains. *Proceedings of the National Academy of Science* 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, 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.

Bermúdez-Otero, Ricardo. Forthcoming. *Stratal Optimality Theory*. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change: Quantity in English and Germanic. PhD Thesis, University of Manchester and Universidad de Santiago de Compostela.

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus* 25.1: 3–103.
Blaho, Sylvia. 2008. The syntax of phonology: A radically substance-free approach. PhD Thesis, University of Tromsø.
Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
Blust, Robert. 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics* 43.1: 73–148.
Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.
Brentari, Diane. 1990. Theoretical foundations of American Sign Language phonology. Doctoral dissertation, University of Chicago.
Broselow, Ellen. 2003. Marginal phonology: Phonotactics on the edge. *The Linguistic Review* 20: 159–193.
Buckley, Eugene. 2002. Rule naturalness and the acquisition of phonology. Handout from a talk given at NAPhC2, Montréal, 16 April 2002.
Calabrese, Andrea. 2005. *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton De Gruyter.
Chomsky, Noam. 1951. Morphophonemics of Modern Hebrew. Master's Thesis, University of Pennsylvania.
Chomsky, Noam. 1955. The logical structure of linguistic theory. PhD Thesis, University of Pennsylvania.
Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
Chomsky, Noam. 1965. *Aspects of the Theory Syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1967. Some general properties of phonological rules. *Language* 43.1: 102–128.
Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacob and Peter S. Rosenbaum, eds., *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz, ed., *Ken Hale: A Life in Language*, pp. 1–52. Cambridge, MA: MIT Press.
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 Morris Halle, Horace Lunt, Hugh McLean and Cornelis van Schooneveld, eds., *For Roman Jakobson: Essays on the Occasion of His Sixtieth Birthday*, pp. 65–80. The Hague: Mouton.
Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–297.
Clements, G.N. 2000. In defense of serialism. *The Linguistic Review* 17: 181–197.
Davis, Stuart and Natsuko Tsujimura. 1991. An autosegmental account of Japanese verbal conjugation. *Journal of Japanese Linguistics* 13: 117–144.
de Lacy, Paul. 2002. The formal expression of markedness. Doctoral dissertation, UMass Amherst.
de Lacy, Paul. 2006a. *Markedness*. Cambridge: Cambridge University Press.
de Lacy, Paul. 2006b. Transmissibility and the role of the phonological component. *Theoretical Linguistics* 32.2: 185–196.
de Lacy, Paul, ed. 2007. *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
de Lacy, Paul and John Kingston. 2006. Synchronic explanation. Manuscript, Rutgers University and UMass Amherst.
Dinnsen, Daniel and Judith Gierut, eds. 2008. *Optimality Theory, Phonological Acquisition and Disorders*. London: Equinox.
Donegan, Patricia and David Stampe. 1979. The study of natural phonology. In Daniel Dinnsen, ed., *Current Approaches to Phonological Theory*, pp. 126–173. Bloomington: Indiana University Press.
Dunn, Carla. 1983. A framework for generalization in disordered phonology. *Journal of Childhood Communication Disorders* 7: 47–58.

Eckman, Fred. 1984. Universals, typologies and interlanguages. In William Rutherford, ed., *Language Universals and Second Language Acquisition*, pp. 79–105. Philadelphia: John Benjamins.

Elbert, Mary, Daniel A. Dinnsen and Thomas W. Powell. 1984. On the prediction of phonologic generalization learning patterns. *Journal of Speech and Hearing Disorders* 29: 309–317.

Féry, Caroline and Ruben van de Vijver, eds. 2003. *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press.

Finley, Sara. 2008. Formal and cognitive restrictions on vowel harmony. Doctoral dissertation, Johns Hopkins University.

Finley, Sara, and William Badecker. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61: 423–437.

Flemming, Edward. *Auditory Representations in Phonology*. New York: Routledge.

Gamble, Geoffrey. 1989. Spanish loans in Wikchamni. In Mary Ritchie Key and Henry M. Hoenigswald, eds., *General and Amerindian Ethnolinguistic in Remembrance of Stanley Newman*, pp. 123–128. Berlin: Mouton de Gruyter.

Gierut, Judith. 1985. On the relationship between phonological knowledge and generalization in misarticulating children. Doctoral dissertation, Indiana University.

Gierut, Judith, Mary Elbert and Daniel Dinnsen. 1987. A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research* 30: 462–479.

Goldsmith, John. 2008. Generative phonology in the late 1940s. *Phonology* 25: 37–59.

Goldsmith, John and Bernard Laks. 2006. Generative phonology: Its origins, its principles, and its successors. To appear in Linda Waugh, John E. Joseph and Monique Monville-Burston, eds., *The Cambridge History of Linguistics*. Cambridge: Cambridge University Press.

Goldsmith, John and Jason Riggle. 2012. Information theoretic approaches to phonological structure: The case of Finnish vowel harmony. *Natural Language and Linguistic Theory* 30: 859–896.

Green, Antony. 2004. Opacity in Tiberian Hebrew: Morphology, not phonology. *ZAS Papers in Linguistics* 37: 37–70.

Hale, Mark and Charles Reiss. 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.

Harris, John and Geoff Lindsey. 1995. The Elements of phonological representation. In Jacques Durand and Francis Katamba, eds., *Frontiers of Phonology*, pp. 34–79. London and New York: Longman.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: The early stages. In René Kager, Joe Pater and Wim Zonneveld, eds., *Fixing Priorities: Constraints in Phonological Acquisition*, pp. 158–203. Cambridge: Cambridge University Press.

Hayes, Bruce and James White. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44.1: 45–75.

Hayes, Bruce and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.3: 379–440.

Hayes, Bruce, Kie Zuraw, Peter Siptár and Zsuzsa Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85: 822–863.

Heinz, Jeffrey. 2008. Phonology I. Manuscript, University of Delaware. Accessed 22 September 2016 from https://goo.gl/KjKpOV.

Hyman, Larry M. 1989. Accent in Bantu: An appraisal. Paper presented at 20th Annual Conference on African Linguistics, April 20, 1989, University of Illinois, Urbana. *Studies in the Linguistic Sciences* 19: 115–134.

Hyman, Larry. 2001. Privative tone in Bantu. In Shigeki Kaji, ed., *Cross-Linguistic Studies of Tonal Phenomena*, pp. 237–257. Tokyo: Institute for the Study of Languages and Cultures.

Idsardi, William. 1997. Sympathy creates chaos. Manuscript, University of Delaware.

Iosad, Pavel. 2012. *Representation and variation in substance-free phonology A case study in Celtic*. Doctoral dissertation, University of Tromsø.

Jakobson, Roman, Gunnar Fant and Morris Halle. 1952. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge: MIT Press.

Jesney, Karen. 2011. Cumulative constraint interaction in phonological acquisition and typology. Doctoral dissertation, University of Massachusetts, Amherst.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kaplan, Aaron. 2011. Variation through markedness suppression. *Phonology* 28: 331–370.

Karttunen, Lauri. 1998. The Proper Treatment of Optimality in Computational Phonology. In *The Proceedings of FSMNLP'98: International Workshop on Finite-State Methods in Natural Language Processing*, June 29–July 1, 1998, pages 1–12, Bilkent University, Ankara, Turkey.

Kaye, Jonathan. 1988. On the interaction of theories of Lexical Phonology and theories of phonological phenomena. In Wolfgang Dressler et al., eds., *Phonologica*, pp. 141–155. Cambridge: Cambridge University Press.

Kaye, Jonathan. 2000. A user's guide to Government Phonology. Manuscript, University of Ulster.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2: 305–328.

Kiparsky, Paul. 1972. Explanation in linguistics. In Stanley Peters, ed., *Goals in Linguistic Theory*, pp. 189–225. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kiparsky, Paul. 1973. Phonological representations. In Osamu Fujimura, ed., *Three Dimensions of Linguistic Theory*, pp. 1–135. Tokyo: TEC Co.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17: 351–367.

Kiparsky, Paul. 2006. Amphichronic linguistics vs. Evolutionary Phonology. *Theoretical Linguistics* 32: 217–236.

Kiparsky, Paul. 2010. Reduplication in Stratal OT. In Linda Uyechi and Lian Hee Wee, eds., *Reality Exploration and Discovery: Pattern Interaction in Language & Life*, pp. 125–142. (Festschrift for K.P. Mohanan). Stanford: CSLI Publications.

Kiparsky, Paul. 2011. Chains or strata? The case of Maltese. Lingbuzz/001379.

Kiparsky, Paul and Carol Kiparsky. 1971. Fact. In Manfred Bierwisch and Karl Erich Heidolph, eds., *Progress in Linguistics*, pp. 143–173. The Hague: Mouton.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.

Kochanski, Greg. 2005. How could one do a statistical test of Optimality Theory? Accessed 30 July 2016 from http://kochanski.org/gpk/teaching/0501Oxford/optimality.pdf.

Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19: 219–251.

LSHSS. 1988 = Special issue (vol. 19, no. 3) of Language, Speech and Hearing Services in the Schools devoted to generalization.

McCarthy, John. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38: 169–195.

McCarthy, John. 1988a. Feature geometry and dependency: A review. *Phonetica* 45: 84–108.

McCarthy, John. 1998b. Sympathy and phonological opacity. Manuscript, University of Massachusetts. ROA-252.

McCarthy, John. 1999a. *Introductory OT on CD-ROM*. Amherst, MA: Graduate Linguistic Students' Association.

McCarthy, John. 1999b. Sympathy and phonological opacity. *Phonology* 16: 331–399.

McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John. 2007. *Hidden Generalizations*. London: Equinox Publishing.

McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, eds., *Papers in Optimality Theory*. University of Massachusetts Occasional Papers 18, pp. 249–384. Amherst, MA: Graduate Linguistic Student Association. [Rutgers Optimality Archive 60, http://roa.rutgers.edu].

McMahon, April. 1992. Lexical phonology and diachrony. In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen and Irma Taavitsainen, eds., *History of Englishes: New Methods and Interpretations in Historical Linguistics*, pp. 167–190. Berlin: de Gruyter.

Meyers, Scott P. 1990. *Tone and the Structure of Words in Shona*. New York: Garland Press.

Miller, Brett, Neil Myler and Bert Vaux. 2016. Phonology in Universal Grammar. In Ian Roberts, ed., *The Oxford Handbook of Universal Grammar*, pp. 153–182. Oxford: Oxford University Press.

Mohanan, K.P. 2000. The theoretical substance of the optimality formalism. *The Linguistic Review* 17.2–4: 143–166.

Moreton, Elliott. 1999. *Non-Computable Functions in Optimality Theory*. Manuscript, UMass-Amherst. Published 2004 in John J. McCarthy, ed., *Optimality Theory in Phonology: A Reader*. Malden, MA: Wiley-Blackwell. Accessed 20 July 2016 from http://roa.rutgers.edu/files/364-1199/roa-364-moreton-1.pdf.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25: 83–127.

Nevins, Andrew. 2007. Review of Andrea Calabrese, Markedness and economy in a derivational model of phonology. *Journal of Linguistics* 10: 223–229.

Nevins, Andrew. 2010. Two case studies in phonological universals: A view from artificial grammars. *Biolinguistics* 4: 217–232.

Nevins, Andrew and David Braun. 2009. The role of underlying representations in L2 Brazilian English. In Andrea Calabrese and Leo Wetzels, eds., *Studies in Loan Phonology*, pp. 181–191. Amsterdam: John Benjamins.

Nevins, Andrew and Bert Vaux. 2007. Underlying representations that do not minimize grammatical violations. In Sylvia Blaho, Patrik Bye and Martin Krämer, eds., *Freedom of Analysis?*, 35–61. Berlin: Mouton de Gruyter.

Newmeyer, Frederick J. 2004. *Possible and Probable Languages*. Oxford: Oxford University Press.

Newton, Brian. 1972. *The Generative Interpretation of Dialect: A Study of Modern Greek Phonology*. Cambridge: Cambridge University Press.

Odden, David. 1980. Associative tone in Shona. *Journal of Linguistic Research* 1.2: 37–51.

Odden, David. 2008. Ordering. In Bert Vaux and Andres Nevins, eds., *Rules, Constraints, and Phonological Phenomena*, pp. 61–120. Oxford: Oxford University Press.

Ohala, John. 1971. The role of physiological and acoustic models in explaining the direction of sound change. *Project on Linguistic Analysis Reports* (Berkeley) 15: 25–40.

Ohala, John. 1972. How to represent natural sound patterns. *Project on Linguistic Analysis* (Berkeley) 16: 40–57.

Ohala, John. 1975. Phonetic explanations for nasal sound patterns. In Charles A. Ferguson, Larry Hyman and John J. Ohala, eds., *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, pp. 289–316. Stanford: Language Universals Project.

Ohala, John. 1981. The listener as a source of sound change. In C. S. Masek, R. Hendrick and M. F. Miller, eds., *Papers from the Parasession on Language and Behavior*, pp. 178–203. Chicago: Chicago Ling. Soc.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In P.F. Neilage, ed., *The Production of Speech*, 189–216. New York: Springer.

Ohala, John J. 1997. Aerodynamics of Phonology. In *Proceedings of the Seoul International Conference on Linguistics*, 20th edition, Seoul: Linguistic Society of Korea, 92–97, http://linguistics.berkeley.edu/~ohala/papers/SEOUL2-aero.pdf (retrieved 24 March 2012).

Ohala, John. 2005. Phonetic explanations for sound patterns: Implications for grammars of competence. In W. J. Hardcastle and J. M. Beck, eds., *A Figure of Speech: A Festschrift for John Laver*, pp. 23–38. London: Erlbaum.

Orgun, Orhan. 1996. Sign-based morphology and phonology, with special attention to Optimality Theory. Doctoral dissertation, University of California-Berkeley.

Paster, Mary. 2013. Rethinking the 'duplication problem'. *Lingua* 126: 78–91.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In Harry van der Hulst, René Kager and Wim Zonneveld, eds., *The Prosody Morphology Interface*, pp. 310–343. Cambridge, UK: Cambridge University Press.

Pater, Joe. 2001. Austronesian nasal substitution revisited: What's wrong with *NC̥ (and what's not). In Linda Lombardi, ed., *Segmental Phonology in Optimality Theory: Constraints and Representations*, pp. 159–182. Cambridge, UK: Cambridge University Press.

Pater, Joe and Anne-Michelle Tessier. 2003. Phonotactic knowledge and the acquisition of alternations. In Maria-Josep Solé, Daniel Recasens and Joaquín Romero, eds., *Proceedings of the 15th International Congress of Phonetic Sciences*, pp. 1177–1180. Barcelona: Universitat Autònoma de Barcelona.

Peperkamp, Sharon, Katrin Skoruppa and Emmanuel Dupoux. 2006. The role of phonetic naturalness in phonological rule acquisition. In David Bamman, Tatiana Magnitskaia and Colleen Zaller, eds., *Proceedings of the 30th Annual Boston University Conference on Language Development*, pp. 464–475. Somerville, MA: Cascadilla Press.

Pesetsky, David. 2013. Comment on Norbert Hornstein, "Evans-Levinson: The sound and the fury". Accessed 29 September 2016 from http://facultyoflanguage.blogspot.co.uk/2013/05/evans-levinson-sound-and-fury.html.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar* (Technical Report). New Brunswick, NJ: Rutgers Center for Cognitive Science.

Pycha, Anne, Pawel Nowak, Eurie Shin and Ryan Shosted. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In Gina Garding and Mimu Tsujimura, eds., *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pp. 423–435. Somerville, MA: Cascadilla Press.

Raimy, Eric. 2000. *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter.

Rasin, Ezer and Roni Katzir 2014. A learnability argument for constraints on underlying representations. Manuscript, MIT accessed on 25 September 2016 from ling.auf.net/lingbuzz/002260/current.pdf.

Rosenthall, Sam. 2006. Glide distribution in Classical Arabic verb stems. *Linguistic Inquiry* 37.3: 405–440.

Saffran, Jenny and Erik Thiessen. 2003. Pattern induction by infant language learners. *Developmental Psychology* 39: 1926–1928.

Samuels, Bridget. 2009. The structure of phonological theory. Doctoral dissertation, Harvard University.

Sayeed, Ollie. 2016. Phonological Time Travel: 'Counterfeeding from the Future'. In *Paper presented at the Linguistics Association of Great Britain annual meeting*, University of York, September 8.

Schane, Sanford, Bernard Tranel and Harlan Lane. 1974/1975. On the psychological reality of a natural rule of syllable structure. *Cognition* 3/4: 484–494.

Scheer, Tobias. 2004. *A Lateral Theory of Phonology. Vol 1: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.

Scheer, Tobias. 2010. What OT is, and what it is not: Review of de Lacy (ed.) 2007. *Journal of Linguistics* 46: 193–218.

Scheer, Tobias. to appear. Optimality Theory (in phonology). Entry in Karlík, Petr and Marek Nekula, eds., Encyklopedický slovník češtiny.

Scheer, Tobias. 2012. *Direct Interface and One-Channel Translation*. Volume 2 of a *Theory of Lateral Phonology*. Berlin: Walter de Gruyter.

Sommerstein, Alan. 1974. On phonotactically motivated rules. *Journal of Linguistics* 10: 71–94.

Spruit, Arie. 1986. Abkhaz studies. Doctoral dissertation, Leiden University.

Stampe, David. 1969. The acquisition of phonetic representation. *CLS* 5: 443–453.

Stampe, David. 1973. A dissertation on natural phonology. Doctoral dissertation, University of Chicago.

Stampe, David. 1979. *A Dissertation on Natural Phonology*. New York: Garland Publishing Co.

Staroverov, Peter. 2010. Too-many-solutions and Reference to Position in Serial OT. *Proceedings of the 33rd Annual Penn Linguistics Colloquium* 16: 205–214.

Stokoe, William. 1960. *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. (Studies in Linguistics, Occasional papers, 8.). 2d ed. Buffalo: Department of Anthropology and Linguistics, University of Buffalo, Silver Spring, MD: Linstok Press, 1978.

Tessier, Anne-Michelle. 2010. Review of Dinnsen and Gierut 2008. *Language* 86.3: 716–720.

Tessier, Anne-Michelle. 2015. *Phonological Acquisition: Child Language and Constraint-Based Grammar*. London: Palgrave-Macmillan.

Tihonova, Olga. 2009. Acquisition and opacity. Master's dissertation, University of Tromsø.

Tyler, Ann A., Mary Louise Edwards and John H. Saxman. 1987. Clinical application of two phonologically based treatment procedures. *Journal of Speech and Hearing Disorders* 52: 393–409.

Uffmann, Christian. 2004. Re-Evaluating Transfer in SLA. *Talk presented at the Montreal Dialogues Workshop*, Université du Québec à Montréal.

Vaux, Bert. 1998. *The Phonology of Armenian*. Oxford University Press.

Vaux, Bert. 2005. Formal and Empirical Arguments for Morpheme Structure Constraints. *Paper presented at the Linguistic Society of America Annual Meeting*, San Francisco.

Vaux, Bert. 2008. Why the phonological component must be serial and rule-based. In Bert Vaux and Andrew Nevins, eds., *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press.

Vaux, Bert. 2016. On the Interactions of Epenthesis and Voice. *Paper presented at the North American Phonology Conference*, Concordia University, Montreal.

Vaux, Bert and Andrew Nevins. 2003. Underdetermination in Language Games: Survey and Analysis of Pig Latin Dialects. *Linguistic Society of America Annual Meeting*, Atlanta.

Vaux, Bert and Samuels, Bridget. 2017. Consonant epenthesis and markedness. In Bridget D. Samuels, ed., *Beyond Markedness in Formal Phonology*, pp. 69–100. Amsterdam: John Benjamins.

Wauquier, Sophie, Tobias Scheer and Marc van Oostendorp. 2015. Projet de Réseau international (GDRI-CNRS): Phonological Theory Agora. Manuscript, Université Nice Sophia Antipolis.

Weigel, William. 2005. Yowlumne in the twentieth century. Doctoral dissertation, University of California-Berkeley.

Whitney, William. 1889. *Sanskrit Grammar*. Cambridge, MA: Harvard University Press.

Wilson, Colin. 2003. Experimental investigation of phonological naturalness. In Gina Garding and Mimu Tsujimura, eds., *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pp. 533–546. Somerville, MA: Cascadilla Press.

Wilson, Colin. 2006. Counterfeeding from the past. Manuscript, University of California, Los Angeles. Accessed 30 July 2016 from http://camba.ucsd.edu/blog/sadphig/files/2009/01/counterfeeding-fromthepast.pdf.

Wiltshire, Caroline. 2003. Beyond codas: Word and phrase-final alignment. In Caroline Féry and Ruben van de Vijver, eds., *The Syllable in Optimality Theory*, 254–268. Cambridge: Cambridge University Press.

Wolf, Matthew Adam. 2008. Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model. Doctoral dissertation: University of Massachusetts, Amherst.

Wolf, Matthew. 2010. On the existence of counterfeeding from the past. *84th LSA Annual Meeting*, Baltimore, January 9. Accessed 30 July 2016 from http://wolf.phonologist.org/OnTheExistenceOfCFFTP-corrected.pdf.

Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Zoll, Cheryl. 1993. Directionless syllabification and Ghosts in Yawelmani. Manuscript, University of California, Berkeley. [ROW-28].

8

The syntax–phonology interface in Rule-Based Phonology

Heather Newell

8.1 Introduction

This chapter focuses on two issues that are important to the construction of an explanatory theory of phonology: modularity and derivation. Modularity encapsulates the proposal that the vocabulary used by each linguistic sub-domain (phonology, syntax, semantics) is module-specific. It is therefore also concerned with how information is passed from one module to another. Here we concentrate on how morphosyntactic information is transferred from the syntax to the phonology. It is argued that syntactic features have no place in the phonological module, and that this is an important factor for distinguishing between theories of phonology and of the morphosyntax–phonology interface. The discussion of modularity then leads us to the question of derivation, or more precisely, of cyclicity. If morphosyntactic information is illicit in the phonological module then the fact that phonological outputs display evidence of cyclic domains must be due either to properties specific to the phonological module, or to the existence of a derivational engine that sends information to the phonology in pieces. I support the latter view here. How this cycling of the computational system works, and how it affects the phonology, is the second topic of the current chapter.

These two issues are integral to constructing a theory of the morphosyntax–phonology interface. They are also two domains of inquiry where traditional, serial, Rule-Based Phonology (RBP) and parallel Optimality Theory (OT), or Constraint-Based Phonology (CBP), have diverged, sometimes significantly. Within OT, which was traditionally entirely parallelist (and still is in many, if not most, sub-camps), the re-introduction of cyclicity to the derivational computation has closed the gap between the two frameworks. Serial OT frameworks such as Lexical Phonology and Morphology Optimality Theory (LPM-OT) (Kiparsky 2000) or Stratal Optimality Theory (SOT) (Kiparsky 2007; Bermúdez-Otero 2014) are fundamentally quite similar to a cyclic, modular, rule-based framework, but there are still some significant differences. Many of these particular differences (ex. strata-internal rule ordering) have been detailed in Vaux (2008) and in the previous two chapters (Purnell, this volume; Vaux & Myler, this volume), and will therefore not be the focus of the current discussion. The recent explosion of modifications to

CBP theories (ex. Stochastic OT, Boersma 1997; Harmonic Serialism, McCarthy 2000; Candidate Chain Theory, McCarthy 2007; in addition to LPM-OT and SOT) makes a general overview of the distinctions between rule-based and constraint-based theories at the interface beyond the scope of this chapter. I therefore focus here on the argument that an understanding of the link between syntactic and phonological derivations is a crucial precursor to developing the correct theory of purely phonological alternations. I emphasize some pertinent distinctions between RBP and SOT as they are currently presented, as they have similar premises regarding the architecture of the grammar and therefore offer an opportune environment in which to discuss detailed nuances of theory-construction and testing. I offer arguments that a serial, cyclic derivation, where affixes are not assigned to strata-particular constraint rankings, best captures the inner-workings of the generative phonological system.

In section 8.2 we will briefly discuss some historical perspectives and theoretical assumptions pertinent to the rest of this chapter. In section 8.3 critical distinguishing issues specific to modularity and derivation germane to the OT–RBT debate are expounded. In these two sections the reader will find extensive references pointing them to more in-depth discussions of the foundational issues. Some relevant current contributions and research will be presented in section 8.4. This main section will discuss aspects of two specific issues relevant to the SOT–RBT debate: (i) whether the delimitation of cycles is accomplished via representational or derivational means, and its link to OT alignment constraints and the Prosodic Hierarchy; and (ii) the relationship, or lack thereof, between the morphological status of affixes and phonological cycles. Section 8.5 will then conclude with an eye to future potential research directions, discussing the implications of the data in section 8.4 for theories of phonology at and below the Prosodic Word.

8.2 Historical perspectives

To situate any discussion of the morphosyntax–phonology interface it is necessary to discuss one's foundational assumptions regarding (i) the morphosyntax (and the morphology and syntax separately if they are taken to be separate modules), (ii) the phonology, and (iii) their manner of interacting. These foundational suppositions cannot be separated from the history of linguistic theorizing that they emerge from, leading to their inclusion in the present section. The presentation below is brief, of necessity, and should not be taken as a definitive list of resources or viewpoints pertinent to the debate, but rather as an overview of the theoretical underpinnings assumed in the following sections.

Proposals in the morphosyntactic literature have vacillated over whether the syntax and morphology are one single or two separate generative engines. Originally compounded (Chomsky & Halle 1968), the emergence of investigations into morphological irregularity vs morphosyntactic regularity inspired their division, instigating the emergence of theories such as Lexical Phonology and Morphology (LPM) (Kiparsky 1982; Mohanan 1986). There it was proposed that a pre-syntactic word-formation module existed wherein affixes and phonological rules were organized into Levels, or Strata (ex. irregular inflection, regular inflection, various types of derivation, compounding). This was taken to explain the link between certain kinds of phonological phenomena and particular constraints on affix positioning, and is the theoretical precursor to SOT, to be discussed in more detail below. Issues with LPM, such as its inability to correctly predict affixal patterning (Fabb 1988), the necessary introduction of a looping mechanism between levels which led to a lack of predictive power

(Kaisse & Shaw 1985), and its overlooking of the parallels between word-internal and syntactic configurations (Marantz 1997), among other arguments, led to theories that reunified the two modules. This chapter assumes a particular version of this reunification, namely a realizational theory of morphosyntax (à la Distributed Morphology (Halle & Marantz 1993, 1994) or Nano-Syntax (Starke 2010)) where the building blocks of the syntactic generative engine are morphemes, and these morphemes consist of abstract feature bundles which are given phonological form post-syntactically.

On the strictly syntactic side, it is a historically robust stance that the derivation proceeds in steps, or cycles (see Boeckx & Grohmann 2007 for a history of the cycle in the syntactic literature). The current theory of cycles, namely Phases, is assumed here. Developed initially by Chomsky (1999, 2001a) (see also Uriagereka 1999), phases are sub-structures that are sent to Spell-Out at both Phonological Form (PF) and Logical Form (LF). Phases have been undergoing theoretical modifications since their inception, and recently proposed variations and modifications touch on their size and mutability (Embick 2010, 2014; Marvin 2002; Adger 2006; Newell 2008; Bobaljik & Wurmbrand 2014; Svenonius 2004; Epstein & Seely 2002; Bošković 2014; among others). Understanding the exact size and properties of syntactic cycles is crucial to the development of any theory of the morphosyntax–phonology interface, but will not be investigated here. It is assumed herein that phases may be large (EP (traditionally v*P), CP, DP, and perhaps PP), or small (triggered by category-defining derivational heads that project vP, nP, and aP). The main focus of this chapter bears on the question of the manner in which the phonological module interfaces with the morphology and syntax at the point of Spell-Out, and what implications this has for current RBP or CBP proposals.

The phonological foundation of the current chapter is one where, in the words of McCarthy (1988: 84), "if the representations are right, then the rules will follow." It is proposed here that the phonological output is organized both sub- and suprasegmentally into a version of feature geometric and syllabic (or CV) representations, respectively. The exact specifics of the organization of these phonological subcomponents are subject to debate. For discussions of feature geometry see Clements (1985); Sagey (1986); McCarthy (1988); Avery & Rice (1989); and Uffmann (2011, and references therein). For discussions of syllabic organization, or the lack thereof, see Bosch (2011 and references therein) and Scheer (2004, and references therein). An important point of debate that will be discussed herein concerns the relative roles of representation (phonological structure) and derivation (procedural computation) to a theory of phonology. Following Newell & Scheer (2007) and Scheer (2010), the existence of prosodic organization above the foot level is questioned (the Prosodic Hierarchy of Selkirk 1984; Nespor & Vogel 1986), but we will occasionally resort to making use of it below for ease of exposition.

This chapter therefore assumes a completely modular linguistic computational system, where the structure generated by the morphosyntax is interpreted cyclically at the interfaces, and where cycles are governed by purely syntactic means. At the interface with PF, morphemes are replaced with vocabulary items, and their forms are underlyingly structured at the melodic level. Underspecified segmental structure and suprasegmental structure are both projected from these phonological underlying representations upon interpretation at PF, and these underlying and surface representations may be subject to the application of phonological rules. Inviolable phonological constraints, in the pre-OT sense (ex. the Obligatory Contour Principle, Leben 1973), may also be active, and are considered a basic part of rule-based phonological models.

8.3 Critical issues and topics

This section focuses on the main issues to be discussed in this chapter, namely *modularity* and *derivation*. Some general aspects of these issues are discussed, and in section 8.4 we delve into some specific subjects that will tease apart certain differing predictions of SOT and the RBP supported here.

Modularity speaks to *what kind of information* is processed in each of the morphosyntactic, semantic, and phonological computational systems. Within a strictly modular framework, the morphosyntactic and phonological systems are non-overlapping. Morphosyntactic information must undergo translation into phonological primes before the phonological system can act upon it. In other words, the syntax manipulates strictly syntactic features and representations, while the phonology manipulates strictly phonological features and representations (Zwicky & Pullum 1986). This entails that syntactic information such as *3rd person* or *XP* is invisible to, or rather, non-existent, in the phonological domain, while a phonological element such as *nasal* or *Onset* can play no part in the syntactic domain. This entailment holds universally. In no language does a person feature trigger, say, lenition, nor does a phonological feature trigger movement.[1] This notion of modularity bears not only on the distinction between phonological and morphosyntactic primes, or features, but also on the question of the translation of cyclic domain edges into phonological terms. It is clear that the syntactic derivation effects the size of the domains for phonological rule application, but how this influence is exerted is open to debate. It was proposed in the post-SPE era that the # and + boundary symbols characteristic of early generative derivations were diacritics (not native phonological objects) and therefore had no place in a theory of phonology. This brought about the rise of indirect-reference theories, the most popular and widespread of which is the Prosodic Hierarchy (Selkirk 1984, 2011; Nespor & Vogel 1986). Recently Scheer (2008) has questioned the existence of the Prosodic Hierarchy as a phonological object, and reopened the debate over the existence of diacritics in the phonology, proposing that structures such as the Prosodic Word (PWd) are merely diacritic. This brings us to an overlapping point pertinent to the spheres of modularity and derivation. Cyclicity in the phonology can potentially be determined in two different ways: derivationally or representationally. A detailed discussion of this distinction can be found in Bermúdez-Otero & Luis (2009); Newell & Scheer (2007); and Scheer (2010). In section 8.4.1 we will examine how syntactically driven cyclic interpretation makes independently verifiable predictions, and will discuss whether the representational cyclicity introduced in the (morpho)phonology via the Prosodic Hierarchy is therefore redundant. Arguments for and against the necessity of both representational and derivational explanations for the cycle are assessed in section 8.4.2. It is argued that a purely derivational phonology is the theoretically cleaner option. This, in turn, calls into question OT constraints like Align that make reference to the Prosodic Hierarchy. It is of note that many current phonological proposals include rules or constraints that make simultaneous reference to both phonological and morphosyntactic information. This mixed-module approach is common within OT and in RBP, but is not specific to, or fundamentally required by, either. In section 8.3.1 we investigate the issues raised by the mixing of modules, and it is demonstrated that this mixing is problematic for an explanatory theory of phonology. A completely modular theory of phonology is proposed to offer superior explanations.

The discussion of the status of derivation within the phonological module is inextricably entangled with the question of modularity in an additional way. Questions of derivation relate to *when information is processed* and *what triggers interpretation.* The first central question regarding phonological derivation is to determine whether phonological strings are processed in cycles or in one fell swoop. As mentioned in section 8.2, for the most part, historically and currently, cyclic derivation has been the standard assumption (with the notable, and ongoing, hiatus taken by traditional parallel OT (McCarthy & Prince 1993b, 2008) and many of its descendants). RBP has traditionally presupposed cyclicity, along with serial application of phonological rules, and these will be taken as basic operations of phonological computation in this chapter. As mentioned above, phonology being processed in cycles leads to the question of whether these cycles are determined phonologically, morphologically (lexically), or syntactically. This brings us to the second question discussed in this chapter relevant to derivation, central only once one accepts that cyclic interpretation is a property of the grammar: what drives phonological cyclicity? In section 8.3.2 we preview the fundamental questions pertinent to how particular phonological cycles are triggered. In section 8.4.3 we carefully examine the case of cyclic phonology in Ojibwe (Newell & Piggott 2014) and will demonstrate that cycles must be determined in the narrow syntax, without reference to morpho-lexical information.

A central concern of this chapter is therefore a meta discussion of the architecture of the grammar assumed generally by some proponents of rule-based and constraint-based systems. Issues with the underlying assumptions of the theories will be raised, and motivations for the return to a derivational, rule-based phonology will be offered. It is crucial to note, however, that neither modularity nor derivation is incompatible with either RBP or CBP. We will therefore focus on a few instances where the underlying assumptions of current theories differ.

8.3.1 Modularity: the domain division

One of the most important distinctions between a classical OT conception of linguistic processing and that of RBP is that the former is not strictly a theory of phonology. Regardless of the trend that has emerged, where phonologists are the strongest supporters of constraint-based grammars, a classical OT grammar is a ranking of grammatical constraints that may also include constraints on morphological and syntactic structures and derivations (Legendre et al. 1993; Aissen 1999; Hayes 2000). A property of standard Optimality Theoretic proposals is therefore that the grammar is not necessarily modular. Morphological and syntactic constraints can, in principle, be ranked both higher and lower than phonological constraints. Also, single constraints in the literature refer simultaneously to both phonological and extra-phonological properties, as in the following constraint that references both linear order and semantic scope:

(1) SCOPE: Make scope transparent on s-structure.

(Jäger 1999)

This non-separation of domains predicts that many non-occurring patterns could emerge, where phonology and syntax are predicted to be able to interact in ways that are unattested. Remember that phonological features never impact whether a verb raises from T to C, for example, just as no formal syntactic feature influences whether Vowel Harmony may occur. Within a non-modular free ranking of constraints, the restriction of languages to the attested patterns is unexpected. Admittedly, the phonological OT literature makes very little reference to narrow-syntactic configurations or features, and a modular OT at that level is not difficult to conceive. Particularly relevant to this chapter is that a modular OT is, in fact, actively promoted within SOT (Bermúdez-Otero 2012). Yet, it is important to note here that the instances where the phonology and morphosyntax do co-refer are an integral part of the OT phonological system, and do nonetheless bleed over into the SOT framework. The most basic non-modular constraints are those like ALIGN (McCarthy & Prince 1993a, 1993b; McCarthy 2003), WRAP (Truckenbrodt 1999), or MATCH (Selkirk 2011). Consider the following:

(2) Match Phrase (Match (Phrase, φ)):
A phrase in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it φ, in phonological representation.
(Selkirk 2011)

This constraint is one of three (the others being specific to words and clauses) proposed by Selkirk to govern the interface between the syntax and the phonology. Now, the translation from syntax to phonology must be effected somehow, but the formulation of the class of alignment constraints underpins a more general issue for the computational system, and specifically for the phonology. The above constraint is not part of a subset of translation operations (dubbed the Translator's Office by Scheer), but rather is interleaved with purely phonological constraints in the phonological system, as in (3).

(3) a.

$_{clause}$[[verb [noun]$_{NP}$]$_{VP}$]$_{clause}$	BinMin (φ, ω)	Match (Phrase, φ)
i. ι(φ(verb φ(noun)φ)φ)ι	*	
☞ ii. ι(φ(verb noun)φ)ι		*

b.

$_{clause}$[[verb[noun adj]$_{NP}$]$_{VP}$]$_{clause}$	BinMin (φ, ω)	Match (Phrase, φ)
☞ i. ι(φ(verb φ(noun adj)φ)φ)ι		
ii. ι(φ(φ(verb noun)φ adj)φ)ι		*

(Selkirk 2011: 447)

The above tableaux demonstrate blatant non-modularity. Not only is the Match Phrase constraint ranked directly in relation to BinMin – a constraint that regulates phonological binarity within a prosodic domain – but also both the input to GEN and the output candidates are a jumble of phonological (φ, ι), morphosyntactic (verb, noun, adj), and syntactic (clause, NP, VP) information. This mix in formal vocabulary strongly entails the presence of other, non-discussed, constraints never listed in the OT syntax–phonology interface tableaux in the literature, namely those preventing the emergence of syntactic features in the phonological output, such as the following:

(4) a. MAX, CLAUSE: Every instance of CLAUSE in the input must emerge in the output.
b. *CLAUSE: No instance of CLAUSE in the input may emerge in the output.

Richness of the Base, combined with the syntactic features in the underlying structures in (3), implies that the non-emergence of syntactic features in the phonological output must be due to universally undominated constraints such as *CLAUSE (it is unclear what the emergence of the feature CLAUSE in a phonological output would look like). Yet, any time a constraint must be proposed to be universally undominated, or a feature in an input must be excluded from all outputs, we run up against the predictions of the general OT framework, wherein constraint ranking is freely mutable. I suggest here that the fact that constraints like *Clause would be inviolable finds a better solution in a strictly modular grammar. No syntactic and morphological features play a role in the phonology, and therefore the tableau in (3) does not represent any possible derivation within the bounds of Universal Grammar. This has implications for the status of the Prosodic Hierarchy, to be discussed below in section 8.4.1 and section 8.4.2.

Note that the above criticism cannot be restricted to CBP derivations, but that a rule in an RBP framework like the one assumed herein must also apply to a phonological string or structure after the translation of syntactic structure into phonological vocabulary has occurred. In other words, the translation of syntactic domains into phonological domains like PWd and PPh (if they are indeed phonological objects) is not considered to be part of the phonology proper in a radically modular theory (Scheer 2012). At the point where phonological rules are applied, reference to the syntax is impossible. Whether this radical stance is correct has implications for any theory of phonology. It is, however, simpler mechanically to eliminate reference to morphosyntax from an RBT theory than from an OT framework where constraints like Align (along with Faithfulness and Markedness constraints) form one of the fundamental constraint classes.

8.3.2 Derivation: the source of cyclicity

As stated above in section 8.2, it is assumed in this chapter that cycles in the phonology are parasitic on a cyclic syntactic generative engine. It follows that syntactic features relevant to cyclic Spell-Out will determine the size and number of phonological domains. A word like *ungrammaticality* in such a system is computed in four cycles, indicated in (5).

(5) $[[\text{un}]_3[[\text{grammatical}]_1\text{ity}]_2]_4$

The status of *un-* as separate phonologically from its base will be discussed further in section 8.4.1, but it is clear from the lack of nasal assimilation evidenced, and its status as a separate (secondary) stress domain, that the phonology has a means to keep this affix separate from its base: *grammaticality*. In Phase-based terms, the set of morphemes {*un-*} comprises a separately processed subset of lexical items in the derivation (a numeration) and therefore constitutes an independent Spell-Out domain. Then we have evidence that *grammaticality* is processed in two cycles, as indicated by the fact that the stress on *grammatical* from cycle 1 carries over into cycle 2. The stress pattern of *grammàticálity* is distinguished from the stress on monomorphemic, single-cycle long words in English like *Lòllapalóoza* where secondary stress generally falls on the first syllable (Pater 2000). We then must postulate a cycle where *un-* and *grammaticality* are linearized, giving us the four cycles in (5).

Within a LPM-OT or SOT account, it is postulated that there are only three possible strata: stem, word, and phrase. Because we limit our discussion in this chapter to elements that are considered to be words we will restrict our discussion to the first two strata. This adherence to a limited number of strata does not preclude a four-cycle analysis of (5), where the stem stratum may iterate (cycles 1 and 2) as may the word (cycles 3 and 4), but this adherence does coincide with two other pertinent propositions (but see section 8.4.3 for a discussion of the proposal in Bermúdez-Otero (2012) that word-level phonology cannot be subject to cyclic reapplication). The first is the proposal that the particular stratum that a phonological string is submitted to is determined by the particular affix(es) added during a cycle. As with its predecessor, LPM, affixes in SOT are affiliated with strata, and this proposal is therefore subject to many of the arguments against LPM in the literature (see references in section 8.2, and the discussion below in section 8.4.3). Yet, an advantage of the proposal that phonological cycles are morphologically determined is that it accounts easily for the reapplication of stem-level rules, like the main-stress rule, in (5) in a way that a phase-based system like the one assumed here appears to have trouble explaining. In a phase-based system without morphological affiliation we expect that only the morphemes present in the first, innermost phase will affect the phonological domain of the root (here *grammar*) (Marantz 2001). Cycle 1 should therefore delimit the domain for main-stress assignment, contrary to the attested output (cf. the less common $[[\textit{grammátical}]_1\textit{ness}]_2$). That this is not the case is clearly caused by an idiosyncratic property of the affix *-ity*; regardless of the morphosyntactic makeup of the structure to which it attaches, it will be interpreted as part of the main-stress domain. Within an SOT framework this falls out directly. In a Phase-based RBP account an extra lexically sensitive process must be posited that may insert an affix into a previously interpreted phonological domain to account for these facts (Lowenstamm 2010). Fortunately, a possible solution to this problem can be proposed, and an independently motivated class of operations (which may be morphologically or phonologically motivated), dubbed Phonological Merger in Newell & Piggott (2014: 353), may be co-opted to account for this behaviour.

(6) Phonological Merger
$[X\ [\ldots\ldots_{PWd}]] \rightarrow [\not{X}\ [..\ X\ \ldots._{PWd}]]$, where X is an affix.

This operation causes an outer affix to incorporate into the phonological domain built on a previous cycle. Two of the clearest cross-linguistic manifestations of this type of operation are infixation (7a), or Selkirk's (1996: 207) internal clitics (7b), where the clitic, which is syntactically separate from the noun to which it cliticizes, emerges as internal to the stress domain that includes the noun.

(7) a. abso*bloody*lutely
b. ù graad 'to the city' (Serbo-Croatian)

Phonological Merger, as seen in (7b), is not restricted to operations that modify linear order, as will be seen again in section 8.4.3 in the discussion of Ojibwe. The stem-level affiliation of *-ity* in (5) may therefore be ascribed to its phonological (dis)position, rather than its morphosyntactic features. Here it would be the case that affixes like *-ity* target the sub-foot structure in the base they attach to, triggering some necessary re-syllabification which would affect the position of stress.[2] In comparison with the behaviour of affixes like *-ity*, section 8.4.3 demonstrates that a single affix in many languages may have multiple phonological behaviours. It will be argued that the behaviour of these 'multi-stratal' affixes is best determined syntactically, rather than morpho-lexically. The notion that morphemes belong to particular strata is therefore not the deciding factor in their phonological behaviour.

This is in part due to the second issue that coincides with the SOT proposal that there are a limited number of phonological strata, namely that morphosyntactic features that correspond to the notion of 'stem' or 'word' do not exist. Volumes such as Julien (2002) and Newell et al. (2017) expound upon the impossibility of determining the status of the term 'word' syntactically. A morphosyntactic notion of 'stem' is even more elusive. As SOT is a modular theory, where no morphosyntactic features (besides those picked out by alignment) are permitted in the phonology, the imposition on the syntax to determine stratal affiliation of affixes is theoretically unmanageable. Within a phase-based account, where the only relevant morphosyntactic features determining phonological cycles are those proposed to be independently necessary to determine syntactic cycles, this problem is not encountered. That said, it is nonetheless the case that different phonological cycles have particular characteristics, and this fact does not fall out of any particular syntactic aspect of phases. Potential directions to take in independently motivating cyclic phonological differences have a long history based in notions such as the Strict Cycle Condition (Kean 1974: 179). Restrictions on the application of phonological rules and the potential ways in which to determine the phonological behaviour of different cycles will be discussed in section 8.5.

This section has presented an overview of the issues that are critical to the discussion of the particular problems to be presented in section 8.4. We will discuss how the particular behaviour of adjuncts has been used to motivate the necessary presence of the Prosodic Hierarchy in the phonological component, and how a derivational account may eliminate this necessity in section 8.4.1 and section 8.4.2. In section 8.4.3 we will examine multi-cyclic affixes in Malayalam and Ojibwe, and will conclude that the cross-linguistic evidence supports the conclusion that affixes are not lexically affiliated with particular phonological strata. Although the morphosyntax–phonology interface predictions of SOT and the RBP framework presented here agree in many aspects, it will be seen that there are domains wherein their implications are non-overlapping.

8.4 Current contributions and research

Here we will examine particular analyses of phonological phenomena for which RBT premises are argued to offer a simpler, more modular solution than that offered within a theory such as SOT. Section 8.4.1 will begin with a discussion of English negative morphemes that has been recurrent in the literature since at least Booij & Rubach (1984). This discussion is pre-empted in Newell (2008) and Scheer (2012), and rests on the proposition that, due to the syntactic governance of cyclicity, words that appear to be structurally indistinct may nonetheless display divergent derivations. This leads to a discussion in section 8.4.2 of data raised in Bermúdez-Otero & Luis (2009) from European Portuguese (EP). Bermúdez-Otero & Luis contend that the EP data argue definitively for the necessity of a representational distinction at the word level. A solution along the lines of Newell & Scheer (2007) and Newell (2008) will be argued to account for this data within a derivational model as well, eliminating the need for a solution that appeals to the Prosodic Hierarchy. Section 8.4.3 will then discuss the notion of strata or levels in SOT, and discuss how the proposition that morphemes are linked to specific strata does not hold when we look at languages where the same morphemes have divergent phono-syntactic behaviour. We will look in detail at the proposal in Newell & Piggott (2014) that there is no possible OT account for hiatus resolution in Ojibwe. This will lead to a short discussion of affixes whose varying phonological behaviour must be due to modular interface effects rather than to stratal affiliation. The conclusion to be reached here is that, for the data at hand, determination of strata or cycles is effected without reference to the identity of particular morphemes, but rather through syntactic computational mechanisms.

8.4.1 The phonological vs syntactic derivation of English un-/in-

There is a well-known phonological distinction between the negative prefixes *in-* and *un-* in English. The nasal consonant in the former assimilates phonologically in Place (and Manner, for sonorants) to a following consonant, while the nasal in the latter does not.

(8) a. i[m]possible b. u[n]balanced
i[n]tolerable u[n]timely
i[ŋ]congruous u[n]kempt
i[l]licit u[n]lovable
i[ɹ]refutable u[n]rounded

Booij & Rubach (1984) propose, within the framework of Lexical Phonology, that the distinction between the two is representational: *in-* is incorporated into the PWd of its base (it is a Level 1 affix), while *un-* projects its own PWd (Level 2).

(9) a. $[\text{in-possible}]_{\text{PWd}}$ b. $[[\text{un}]_{\text{PWd}}[\text{balanced}]_{\text{PWd}}]$

Assuming that assimilation occurs only within a PWd, the distinction in (8) is captured.
It is quite simple to translate the above into an SOT framework. Here, a constraint requiring assimilation must be ranked higher than a constraint requiring Faithfulness to Place features in the stem stratum (Level 1), where *in-* is introduced (10), while the opposite ranking must be true at the word stratum (Level 2), where *un-* is inserted (11).

(10)

in-possible	ASSIMILATE	$\text{FAITH}_{\text{PLACE}}$
inpossible	*	
☞ impossible		*

(11)

un-balanced	$\text{FAITH}_{\text{PLACE}}$	ASSIMILATE
umbalanced	*	
☞ unbalanced		*

The above, like the representational/LP account, successfully captures the phonological pattern observed. It does, however, leave some questions unanswered. The first is whether the above (re-)rankings are indicative of a larger, cross-linguistic pattern. It has been observed since the beginning of the generative phonological enterprise that the output of a first cycle tends to be preserved on a subsequent cycle (ex. Kean 1974). In versions of OT that countenance multiple constraint rankings/levels this has been translated as the proposal that reranking across strata can only lead to the promotion of Faithfulness constraints in relation to Markedness constraints.

(12) Constraint typology and the limits of re-ranking: The core-periphery organization of the lexicon is the consequence of the fact that, in the typical case, re-ranking is limited to Faithfulness constraints (PARSE and FILL), within an otherwise invariant constraint system.

(Itô & Mester 1995: 183)

The above is not, however, assumed in current OT analyses generally. Bermúdez-Otero & Trousdale (2012), alternately, propose that constraint re-ranking is due to the life cycle of phonetico-phonological processes, leading to the diachronic demotion of Faithfulness constraints in smaller domains. In either case, the distinction between *un-* and *in-* is not linked reliably to a tenet of the OT system.

This brings us to the second question of whether this re-ranking is purely lexico/phonological, or whether the differing levels of Faithfulness across cycles is linked to the synchronic morphosyntactic derivation of particular constructions. In Newell (2005a, 2005b) it is demonstrated that the prefixes *in-* and *un-* not only have distinct phonological behaviour, but also diverge morphosyntactically.

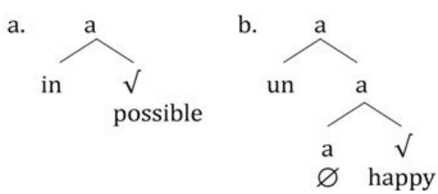
(13)

	(a) phonology	(b) morphosyntax: features	(c) morphosyntax: structure
in-	assimilation	projection of adjectival features	no participation in bracketing paradoxes
un-	no assimilation	no projection of features	participation in bracketing paradoxes

(13a) has already been demonstrated in (8). (13b) is demonstrated by the fact that all words prefixed with *in-* are adjectives, while words prefixed with *un-* may be adjectives, nouns, or verbs.

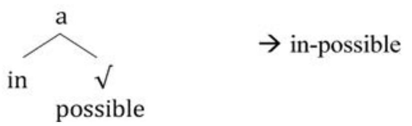
(14) a. inept, impossible, *intie, *imbirthday
b. unhappy, undo, unbirthday[3]

What (14) indicates is a morphosyntactic distinction between the two affixes. Either *in-* selects for an adjectival base, or *in-* is projecting adjectival features. Within a theory such as Distributed Morphology each featureless root morpheme merges (directly or indirectly) with a category-defining head (but see Wiltschko & Déchaine 2010; Borer 2013 for alternate views of lexicalization and the syntax–phonology interface). This proposal, paired with the fact that *in-* and not *un-* attaches to bound roots (ex. *inept*, with the notable exception of *unkempt*), as well as the fact that *in-* emerges in the same phonological domain/cycle as its base, argues for the latter, seen in (15a). The fact that *in-* cannot attach to nouns and verbs indicates that this morpheme may only select for roots or adjectives. As *un-* does not attach to bound roots, and never affects the category of the base to which it attaches, we can conclude that *un-*, as is characteristic of syntactic adjuncts, projects no category features (15b).

(15)


Adjuncts, which do not project morphosyntactic features up the tree, have been argued in the syntactic literature to (i) merge a-cyclically, counter to the Extension Condition of Chomsky (2001b) (Lebeaux 1988; Nissenbaum 2000; Stepanov 2001; Ochi 1999); (ii) be interpreted phonologically prior to their syntactic merger (Uriagereka 1999); and (iii) participate in bracketing paradoxes (Pesetsky 1985; Nissenbaum 2000; Newell 2005a, 2005b, 2008). The divergent syntactic behaviour of adjuncts and non-adjuncts leads to the phonological distinction we see in (8) without calling for additional machinery like prosodic or stratal affiliation in the phonological domain. The derivation for (15a) occurs in one cycle, or phase, *in-* being an adjectival phase head:

(16)



Assuming for the moment that the output of cyclic interpretation to the phonology includes the projection of prosodic structure, the PF output of (15a) will be a single domain. Let us assume this domain to be a PWd. Place assimilation takes place at PF interpretation.

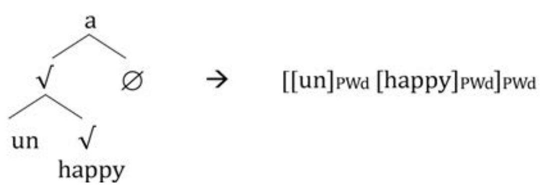
(17) $[\text{impossible}]_{\text{PWd}}$

The derivation of (15b), however, is computed in three steps, the first two of which, given the adjunct nature of *un-*, may occur in parallel. A null adjectival categorizing head triggers the phase in (18a). In (18b) *un-* is the sole member of its numeration/cycle. (18c) is the structure after the a-cyclic merger of *un-*, which must be interpreted at PF to ensure proper linearization of its pieces.

(18) a.

b. un → $[\text{un}]_{\text{PWd}}$

c.



As each PF interpretation leads to prosodic projection, the output of (18) will be as follows the arrow in (18c). Note that the Strict-Layer Hypothesis, a tenet of classic Prosodic Phonology (Selkirk 1981, 1984; Nespor & Vogel 1986), has been abandoned, and therefore the nested PWd structures seen here are permitted by the theory (following Selkirk 2011; Itô & Mester 2013).

The specifics of the phonological rule responsible for assimilation will be expanded on below. What is important here is that the morphosyntactic distinctions between the two negative prefixes in English, along with a cyclic account of syntactic interpretation, leads to the distinction in phonological structure seen in (17) and (18c). This is the exact structural distinction that was proposed in Booij & Rubach (1984). The advantage of the present procedural derivation is that the distinction between *in-* and *un-* is non-stipulative. Syntax feeds phonology. Assimilation here applies within a PWd, and the morphosyntax of *un-* results in its phonological separateness from the base to which it attaches.

Before moving on to the next dataset, it is important to remember the discussion in section 8.3.1, above. The PWd structure assumed in the above derivations is in line with the standardly held view in the phonological literature that the Prosodic Hierarchy is the (indirect) phonological link to syntactic structure. The proposal that the assimilation rule in English is sensitive to this prosodic structure is therefore also standard.

(19) $\text{N} \rightarrow [\alpha\text{Place}]\ /\ \_\_\_\_\ C_{[\alpha\text{Place}]}$ $\quad$ ($^{*}\text{N} \rightarrow [\alpha\text{Place}]\ /\ \_\_\_\_]_{\text{PWd}}\ C_{[\alpha\text{Place}]}$)

The rule in (19) will not apply if a PWd boundary intervenes between the nasal and the following consonant, as the environment for assimilation (direct linear adjacency) will then not be met. This type of explanation is an illustration of a representational account of the phonological derivation (Newell & Scheer 2007; Scheer 2010, 2012). But, the derivations in (16–18) allow for an alternate, purely procedural, account of the manner in which phonological rules are applied. If we assume that the assimilation rule must be applied at the point in the derivation where the nasal consonant is first interpreted, then reference to prosodic structure becomes superfluous. Let us assume that the nasals in both affixes are underspecified for Place. In (16) the nasal is interpreted in the first cycle at PF at the same time as the root *possible* is interpreted. The nasal and the conditioning segment for assimilation are therefore processed in the same cycle, and assimilation may occur. In (18), however, *un-* is first interpreted alone in (18b). As there is no conditioner for assimilation at this point in the derivation, the nasal will emerge with default Place features, here coronal. The proposal here is that these underspecified nasal segments in the input must be fully specified upon interpretation at PF, either by the application of phonological rules or by default projection of features. At a later point in the derivation assimilation will not apply (18c), as the nasal is no longer underspecified. Given this account of the distinction in assimilation patterns here, reference to the PWd is not necessary and therefore not the optimal tool to explain the nasal assimilation patterns. If similar procedural accounts of phonological derivations can be motivated cross-linguistically this will support the proposal of Scheer (2008) that the Prosodic Hierarchy is not the correct way to account for phonological cyclic effects. This type of analysis is discussed further in section 8.5.

The conclusion here is that a rule-based underspecification account of nasal assimilation in English follows from a Phase-based, fully modular account of the derivation. This accounts for both the morphosyntactic and phonological characteristics of these derivations without recourse to either different strata-specific phonologies or diacritic lexical affiliation. A framework like SOT, where Richness of the Base holds only at the input (Bermúdez-Otero 2007), can also account for this, but it is of note that the RBP proposal need not appeal to different constraint rankings/rules to achieve the correct results, making it a simpler account.

8.4.2 European Portuguese and diminutive adjunction

In the above section we managed to eliminate the need to appeal to the Prosodic Hierarchy in our phonological rules while capturing the apparent level distinctions between two prefixes. Bermúdez-Otero & Luis (2009), working with an SOT framework, argue that the phonological cycles in European Portuguese (EP) not only support a stratal account of the phono-syntactic derivation, but also offer crucial evidence for the existence of the Prosodic Hierarchy. It is of note here that these authors are explicitly aware of the potential duplication of explanations, representational vs procedural, exemplified by the discussion of the negative prefixes above and that they attempt to push the possibility of a procedural account as far as possible within the boundaries of SOT. Within SOT, remember, there are a finite number of levels countenanced: stem, word, and phrase (as opposed to the word, phrase, and clause domains of Match Theory (Selkirk 2011)). In this section we offer an alternative account that again appeals to syntactic adjunction and supports an account of the data that need not take recourse to representational domain distinctions.

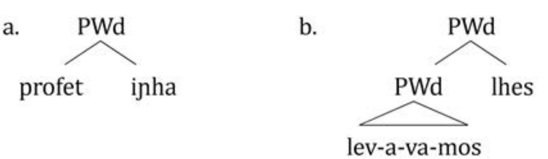
Interestingly, Bermúdez-Otero & Luis show EP affixes/clitics to have (at least) four different behavioural patterns. Characteristic behaviour of suffixes in EP is that they form a single domain for stress assignment with the root, and, if a suffix begins with a high front segment, it will trigger lenition of certain root-final /t/, /k/, or /g/s. Suffixes are therefore clearly stem-level affixes according to Bermúdez-Otero & Luis, or first-Phase affixes within a cyclic RBP account. Proclitics in EP can be separated from the verb by other words, and can scope over conjoined verb phrases. They are clearly separate words from the verb, as opposed to prefixes, and therefore must be introduced at the phrasal stratum within an SOT framework. Enclitics, on the other hand, are not separable from the verb and cannot scope over coordination, so are clearly affixal. But, enclitics never affect the position of main stress in the word, and never trigger root-final lenition. It therefore appears that these enclitics are members of the word stratum. The problem here is the behaviour of the diminutive suffixes *-inho* and *-ito*, which fall between that of the proposed stem and word-level morphemes. The diminutive affixes may be stressed, like stem-level affixes, but do not trigger lenition, like enclitics.

(20) a. profet-ia → profecía (cf. proféta)
‘prophecy’
b. profet-inha → profet-ínha
‘little prophet’

It is therefore not possible for the above patterns to emerge within a purely procedural SOT system.

(21) a. [proclitics [[[root-suffixes] diminutives] enclitics]]
[phrase [[[stem] word?] word?]]
(Bermúdez-Otero & Luis 2009)

If SOT only allows for three strata, then how can the four-way distinction in affix/clitic behaviour be accounted for? The solution proposed by Bermúdez-Otero & Luis is that diminutive morphemes in the language are representationally distinct from enclitics. Both diminutives and enclitics are introduced at the word stratum, but while the diminutives form part of the PWd with the stem affixes, the enclitics are phonologically adjoined to the PWd.

(22) 

The authors therefore conclude that the Prosodic Hierarchy is empirically motivated, as it is the sole way in which to capture this four-way distinction in phonological behaviour within a single language. In other words, their conclusion is that the phonology–syntax interface must be moderated both procedurally (stratally) and representationally (prosodically). The authors admit that this conclusion is linked to the SOT proposal that the strata be limited in number to three. If four levels were allowed for, we would not need to propose a representational distinction between the diminutives and the enclitics, allowing for a purely procedural account of the interface.

Each of these proposals, where SOT is either augmented representationally (by the Prosodic Hierarchy) or procedurally (by adding an extra stratum) is, however, equally stipulative. A fourth stratum, like the representational distinction in (22), is motivated only by the need to account for the four-way distinction in (21). Bermúdez-Otero & Luis, and proponents of SOT in general, limit the number of strata to three in order to constrain the predictive power of the theory (among other motivations). It is unclear, however, how allowing for lexically specified distinctions in representational prosodic structure does not proliferate the number of cycles in exactly the way restricting SOT to three levels is meant to prevent. But, if a proliferation of strata were to be allowed instead, we would be faced with the problem that plagued SOT's theoretical predecessor, LPM, discussed in section 8.3. LMP had to resort to adding levels, and loops between levels, to explain the many phonological distinctions that needed to be accounted for, both within and across languages (Kiparsky 1982; Mohanan 1986). These issues lead us to look for a deeper explanation of the distinctions seen. The data in EP are indicative of a distinction that is real, just like in the case of the English negative prefixes. We can therefore ask ourselves if there is an analogous phono-syntactic account of the affixal divisions in (21).

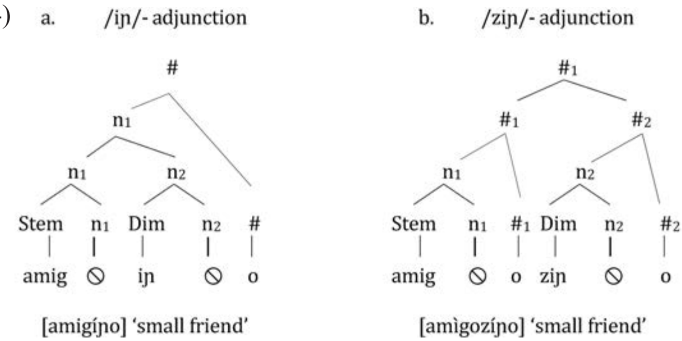
To this end, consider the analysis of the morphosyntax of the diminutive affixes in Brazilian Portuguese. It is argued in Bachrach & Wagner (2007) that the effect diminutives have on the position of stress in a word differs from that of other derivational affixes. Bachrach & Wagner propose that diminutive affixes behave like members of co-compounds (dvandva) rather than like derivational affixes. It is argued that they constitute a second domain for stress, rather than constituting part of the stress domain of the root. This is supported by the interaction of diminutive affixation with mid-vowel raising in unstressed positions (as noted and referenced by Bermúdez-Otero & Luis).

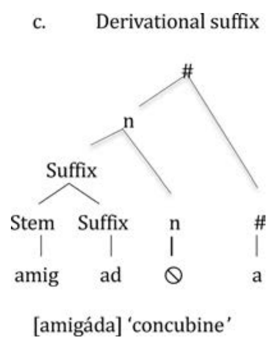
(23) a. bɛ́lo → bel-éz-a ‘beautiful, beauty’
b. bɔ́la → bɔl -íɲ-a ‘ball, (small) ball’
c. bɔ́la → bɔ̀la-zìɲ-a ‘ball, (small) ball’

(Bachrach & Wagner 2007: 2)

Non-diminutive derivational suffixes shift stress away from the root vowel, leading to its raising (23a). The diminutive affix may cause destressing of the root vowel (23b) or demotion of root stress to secondary (23c). In (23b) the root vowel is argued to have been stressed on a previous cycle, and stress clash to have been resolved through deletion. Interestingly, stress clash is not resolved in all dialects (ex. *cafèzínho* Cegalla 2008: 38). In both (23a) and (23b) it is apparent that the derivation includes two cycles of phonological interpretation, where the root vowels receive stress in the first. Why should these diminutive affixes constitute separate cyclic domains from their hosts? According to Bachrach & Wagner, diminutive affixes generally are adjuncts, and are composed of a diminutive root and a category-defining head. Non-adjunct suffixes are proposed to be, on the other hand, monomorphemic.

(24)





(Bachrach & Wagner 2007: 5)

The positions of adjunction in (24a,b) are argued to explain other agreement and distributional facts regarding the diminutive affixes that would take us too far afield from the discussion at hand. For our purposes it is sufficient to note that the adjunct status of the diminutive affixes brings their derivational status in line with that of *un-*; they will be interpreted phonologically prior to merger with their base. The phonological structure of the words in (24) therefore parallels the structures in (16) and (18c), but now this structure has morphosyntactic motivation. The distinction here is that, contrary to Bermúdez-Otero & Luis' analysis, it is the diminutive affixes that are 'PWd' adjoiners.

(25) a. $[[\text{amig}]_{PWd}\,[\text{íɲo}]_{PWd}]_{PWd}$
b. $[[\text{amìgo}]_{PWd}\,[\text{zíɲo}]_{PWd}]_{PWd}$
c. $[\text{amigáda}]_{PWd}$

If the diminutive morphemes are adjuncts to the PWd, as well as being adjuncts in the syntax, then we must ask ourselves what the derivational/phonological properties of the clitics are that cause them to be phonologically farther from their base. The syntax of Portuguese clitics and their relation to phonological Spell-Out will not be solved here, but will undoubtedly be tied to the functional nature of these morphemes and the resistance functional heads show to bearing stress. In any case, it was the distinction between the two 'word-level' suffixes (diminutives and enclitics) that was the motivation for the proposal that the Prosodic Hierarchy is necessary to account for the EP data. Assuming that EP diminutives are adjuncts, as in BP, this entailment no longer holds. Phonological rules apply at PF interpretation, and adjuncts are interpreted prior to their incorporation into the Narrow Syntactic structure. The enclitics, as suffixes, may be interpreted after merger to their base (see (22b)). In this type of structure the enclitic will not bring its own stress into the derivation, and stress determined on the first cycle of PF interpretation will persist.

The entailment is then that it is the timing of interpretation determined by the syntactic structure that distinguishes the problematic levels seen in (21) which, consequently the surface stress patterns in EP, and we can again do away with any reference to the Prosodic Hierarchy. EP suffixes will be interpreted at the same time as the root to which they attach, while enclitics merge to a domain that has already been stressed, and syllabified. The final consonants that do not undergo lenition in the presence of enclitics are therefore structurally distinct from the consonants that do undergo lenition under suffixation, just as in the case of nasal assimilation in English. In the case of diminutives, they themselves constitute separate phases from the base to which they attach. They therefore, as discussed by Bachrach & Wagner, have the phonological behaviour of a kind of compound. Their syntactic structure predicts their phonological differentiation from the enclitics. A fully procedural account is mechanically simpler than the Strata+PH proposal, and accounts for both the morphosyntactic and phonological behaviour of the morphemes involved.

This section has argued that a purely procedural account of phonological domains in EP can be countenanced, leading to a more explanatory account of data that have appeared to require a representational account. It has not been shown, however, that an SOT account cannot account for the data at hand. An SOT framework could be adapted to the account above. The following section, however, elaborates a pattern that has been argued in Newell & Piggott (2014) to be the result of cyclic rule application, and, importantly, to not be derivable in a parallel CBP framework.

8.4.3 Multi-cyclic affixes

It is not uncommon cross-linguistically for a single affix to display different phonological behaviour depending on the construction in which it emerges. An example of this is the causative morpheme in Malayalam (Michaels 2009). The same morpheme, *-ikk*, marks both low and high (traditionally lexical and syntactic) causative constructions. The low causative morpheme either coalesces with the root-final consonant (26a), or resolves an emergent hiatus by deletion of a vowel (26b), while the high causative morpheme does not coalesce with a root-final consonant (27a), and resolves hiatus through epenthesis (27b).

(26)	a.	/aaʈ	+ ikk/	[aaʈʈ]	‘Y shakes X’
		shake	+ cause		
	b.	/nana	+ ikk/	[nanakk]	‘Y waters X’
		water	+ cause		
(27)	a.	/paaʈ	+ ikk/	[paaʈikk]	‘Y makes X sing’
		sing	+ cause		
	b.	/kaɹa	+ ikk/	[kaɹajikk]	‘Y makes X cry’
		cry	+ cause		

This type of pattern is crucially problematic for a non-Stratal Parallelist model such as classical OT, as the input phonological strings in the (a) and (b) examples are indistinguishable in all relevant respects. A non-modular solution to this problem would be to reference the different syntactic structures of (26) and (27) in the constraint ranking (root-attached and non-root-attached causative morphemes, respectively), but this then opens up the issue discussed in section 8.3, where the mixing of syntactic and phonological constraints predicts a sensitivity to the phonology by the syntax that is unattested at the segmental level. To account for the above pattern in SOT, the affix *-ikk* must be permitted to be a member of more than one stratum. This is problematic in that strata are then underdetermined by the morphology, calling into question the premise that strata are morphologically driven (affixes being either stem or word level). Although SOT is a realizational theory of phonology, where cycles are determined in the syntactic structure, it is nonetheless proposed, as seen above, that a single affix at different positions will trigger the cycle to which it is affiliated, as seen with the behaviour of *-ity* discussed in section 8.3.2 (Bermúdez-Otero 2014; but cf. Bermúdez-Otero 2015). Take the example of the affix *-al* in English, as discussed by Bermúdez-Otero (2014). *-al* displays stem-level, stress-affecting behaviour regardless of whether it is affixed to a root (28a) or to a word-level affix (28b).[4]

(28) a. affíxal (cf. áffix)
b. gòvernméntal (cf. góvernment)

Comparing the behaviours of *-al* and *-ikk* it is apparent that their different behaviour must have different sources, and therefore cannot both be due to stratal affiliation being linked to particular syntactic positions. That being the case, it is simpler to propose that the different phonological behaviours of the causative morpheme are determined by non-lexical means, and that it is not the morphological affiliation of these affixes that determines their phonological behaviour. Note that the pattern seen in (26) and (27) is not an isolated case. Productive multi-stratal affixes are found cross-linguistically in unrelated languages such as Malagasy, Acholi, Berber, and Ojibwe (Newell 2014). In addition, Bermúdez-Otero (2013: 12) proposes that stem-level phonology is non-analytical and therefore liable to idiosyncratic behaviour, yet it is clearly not the case that all multi-level affixes display quirky behaviour within the first phonological cycle. Storing predictable phonological forms is arguably undesirable (see also Embick & Halle 2005; Myler 2015 for arguments against stem-level storage).

Setting these problems aside for the moment, a technical account of the Malayalam is possible in SOT, if the causative morphemes are treated as accidentally homophonous (a proposition that is also problematic, as homophony of this type is argued to be dispreferred or impossible by Leu 2015). Although Kilborne-Ceron et al. (2016) argue that both instances of the causative morpheme are heads of an Event Phrase (following Travis 2010), it is possible that these syntactic positions are distinguishable in a way that triggers interpretation at either the stem or word level. Here the domain defined by the low causative morpheme (E_1P) is subject to a stem-level constraint ranking (29), and the domain defined by the high causative (E_2P) to a word-level ranking (30).

(29)

nana+ikk(E_1)	*HIATUS	DEP[J]	FAITH$_V$
nanajikk		*	
☞ nanakk			*
nanaikk	*		

(30)

kaɹa-ikk(E_2)	*HIATUS	FAITH$_V$	DEP[J]
☞ kaɹajikk			*
kaɹakk		*	
kaɹaikk	*		

If all multi-stratal affix behaviour could be accounted for in this way, then the question of whether RBP and CBP are equivalently capable of bridging the interface between syntax and phonology would be indeterminate in this sphere.

The following Ojibwe data demonstrate, however, that this equivalency does not universally hold. Newell & Piggott (2014) argue that the Ojibwe facts are underivable in SOT, and therefore we will expand on their presentation here. Hiatus in Ojibwe (Algonquian) is resolved by deletion within the Event Phrase (EP) in a verbal construction (as seen above for Malayalam) or within nP in a nominal construction (31). Between Tense and EP, hiatus is not resolved if the prefix is bi-moraic, and is resolved by epenthesis if the prefix is monomoraic (32). Between a person prefix and its base, hiatus is resolved by epenthesis in verbal and alienable possession constructions (33a,b), but by deletion in inalienable possession constructions (33c).

(31) Hiatus resolution in the EP domain

a. [nigi:[wa:biwe:ʒi:na:na:nig]$_{EP}$]$_{CP}$ ‘we painted them white’
ni-gi:-wa:bi-we:ʒi:-in-a:-ina:ni-Ø-ag
‘1-past-white-paint-final-TS(3 theme)-1plural-Ind-3plural’

b. [o[name:miwa:n]$_{nP}$]$_{DP}$ ‘their sturgeon(s)’
o-name:-im-(i)wa:-an
‘3-sturgeon-possessive-3plural-obviative’

(32) Hiatus resolution between Tense and EP

a. [gi:[a:gamose:]$_{EP}$]$_{CP}$ ‘he walked in snowshoes’
gi:-a:gam-ose:
‘past-snowshoe-walk’

b. [nigà[dá:gamòsè:]$_{EP}$]$_{CP}$ ‘I will (probably) walk in snowshoes’
ni-ga-a:gam-ose:-Ø
‘1-future-snowshoe-walk-Fin’

(33) Hiatus resolution between Person Marker and its base

a. [ni[da:gamose:$_{EP}$]$_{CP}$ ‘I walk in snowshoes’
ni-a:gam-ose:
‘1-snowshoe-walk’

b. [ni[dakwe:m $_{nP}$]$_{DP}$ ‘my wife’
ni-akwe:-im
‘1-woman-possessive’

c. [no:komis]$_{DP}$ ‘my grandmother’
ni-o:komis
‘1-grandmother’

Given the fact that the data in (31–33) consist uniquely of words, and not phrases, let us again restrict discussion of an SOT account to two strata (stem, word). As the suffixes in

(31) emerge in syntactic positions that are closer to the root than the prefixes (inside EP and nP), hiatus resolution by deletion must be effected at the stem level, as seen below for the first cycle of (31b).

(34)

name:-im-(i)wa:-an	*HIATUS	DEP[d]	FAITH_V
☞ name:miwa:n			**
name:imiwa:an	**		
name:dimiwa:dan		**	

A re-ranking of constraints at the word level will give the correct output for non-resolution between the tense morphemes as in (32a), here in (35). An additional restriction on monomoraic prefixes such as (36a) will force epenthesis in cases like (32b (and 33a,b)), seen in (36b).

(35)

gi:-a:gam-ose:	FAITH_V	DEP[d]	*HIATUS
☞ gi:a:gamose:			*
gi:gamose:	*		
gi:da:gamose:		*	

(36) a. *HIATUS(μ):
A monomoraic vowel may not be followed by a vowel.

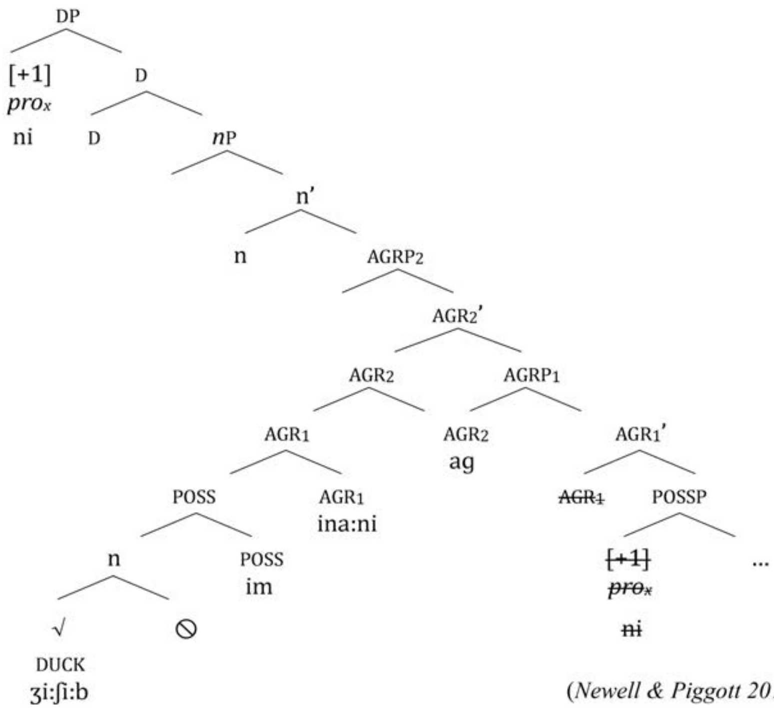
b.

ni-ga-a:gam-ose:	FAITH_V	*HIATUS(μ)	DEP[d]	*HIATUS
nigaa:gamose:		*		*
niga:gam-ose:	*			
☞ nigada:gamose:			*	

The problem arises when trying to account for the distinct hiatus resolution strategies in (33b) and (33c). Unlike in the case of Malayalam causatives, the person prefixes in the Ojibwe possession constructions cannot be stacked. It is argued in Newell & Piggott (2014) that they emerge in DP: in the identical syntactic position in each construction. Relevant structures for alienable and inalienable derivations are seen in (37) and (38), respectively (strikethroughs indicate elements that have moved).

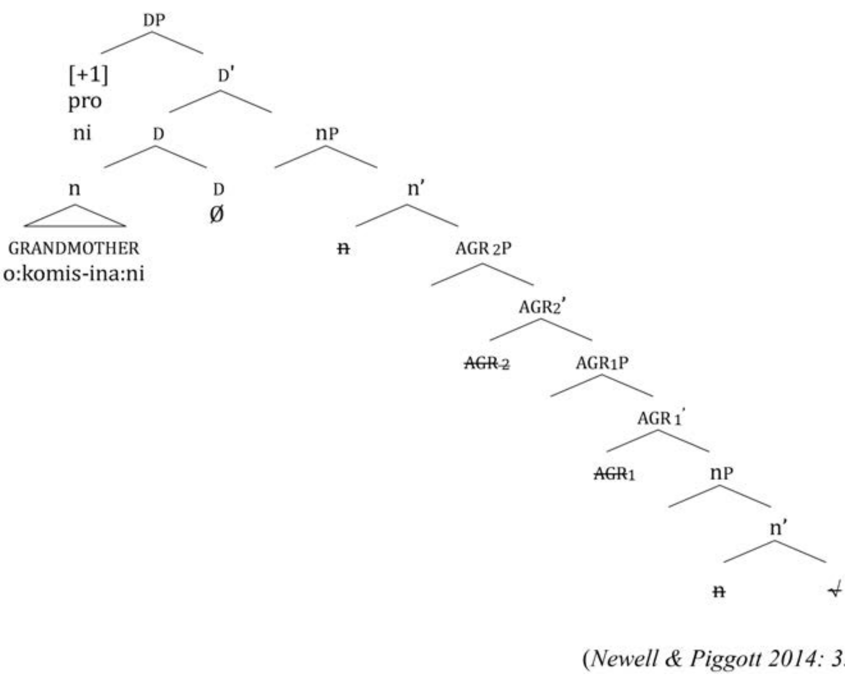
It is therefore impossible to distinguish the two prefixes based on either their featural or surface distributional properties. Newell & Piggott argue that the derivation in (33c) seen in (38) has the attested output due to the fact that in inalienable possession construction the root (here *o:komis*) raises to D to check its argument structure features (inalienable nouns, unlike alienable nouns, are ineffable in Ojibwe without a possessive prefix). The derivation in (37) shows that the root, here *ʒi:ʃi:b*, does not raise out of nP. The root and its suffixes are therefore interpreted in the first phase (nP), while the prefix is not introduced or interpreted until the second phase (DP), parallel to the EP enclitic derivation in the previous section.

(37)



(*Newell & Piggott 2014: 350*)

(38)



(*Newell & Piggott 2014: 350*)

In (38) both *ni-* and *o:komis* are interpreted in the same cycle, as they emerge in the same phase (DP). It is this fact, rather than any fact about the particular properties of the person prefix, that determines that hiatus will be resolved through deletion. Note that the person prefixes are not amenable to an account like that for *-al*, as they do not always behave like stem-level affixes (in fact, stem-level behaviour is exceptional for these affixes), nor are they amenable to a multi-level account, as their syntactic positions upon PF interpretation are invariable. The Ojibwe derivations therefore evidence multiple issues for a stratal, constraint-based account. First, as seen above, morphological affiliation to a particular stratum cannot account for the behaviour of the person prefix. Secondly, non-analytic listing of stem-level expressions would force the storage of a large number of predictable inflected forms in the language, a proposal that is clearly undesirable.

Furthermore, Ojibwe shows cyclic reapplication of processes that must be considered word level in SOT, a phenomenon that is proposed to be unattested in Bermúdez-Otero (2012): section 8.3.1. If both the Tense and person prefixes are word-level affixes, and the word level cannot iterate, this predicts that either word-level phonology will only apply to the innermost word-level affix, or that word-level phonology will apply to both prefixes within one application of the cycle. Neither of these derivations gives the correct result. Newell & Piggott propose that there is a more nuanced explanation for why epenthesis occurs than stated in the constraint in (36a). It is proposed therein that degenerate feet are illicit at the left edge of a PWd. A monomoraic prefix will therefore move into the PWd to its right to permit the construction of a licit prosodic structure, another instance of Phonological Merger, repeated here.

(39) Phonological Merger
$[X\ [\ .\ .\ .\ .\ .\ .\ _{PWd}]] \rightarrow [X\ [.\ .\ X\ .\ .\ .\ .\ _{PWd}]]$, where X is an affix.

In line with the attempt in this chapter to avoid a dependency on the upper levels of the Prosodic Hierarchy, note that these prefixes cannot project a licit foot. This alone can motivate their incorporation into the domain to their right, erasing the need to reference the PWd in (39). Consider (40): a reiteration of (32b) showing foot structure and a more detailed syntactic bracketing, which contains both a tense and a person prefix.

(40) [(nigà)(dá:)(gamò)(sè:)]$_{PWd}$ 'I will (probably) walk in snowshoes'
[ni[-ga-[a:gam-ose:-Ø]$_{EP}$]$_{TP}$]$_{CP}$
'1-future-snowshoe-walk-Fin'

Two things are pertinent here. First, note that hiatus is resolved by epenthesis between the tense morpheme and the verb. This indicates that Phonological Merger has occurred. Secondly, note that the person prefix and tense morphemes are footed together, as indicated by the secondary stress on *ga-* and not on *ni-*. Ojibwe stress is exhaustive, and degenerate feet are permitted as a last resort, at the right edge of a domain. It follows that the phonology of the word in (40) is computed in three cycles. The first is the PF interpretation of the EP: [*(a:)(gamo)(se:)*]. If hiatus emerged in this domain it would be resolved by deletion, as in (31). The next cycle of interpretation is the complement of C: TP. Here *ga-*, being monomoraic, undergoes Phonological Merger into the domain projected on cycle 1, giving ~~ga~~[*(gada:)(gamo)(se:)*]. Here hiatus is resolved through epenthesis. Note that it cannot be the case that *ni-* and *ga-* are interpreted in the same cycle. If they were, they could be footed together at Spell-Out, and we would predict that they, like the bi-moraic tense prefixes, would neither undergo Phonological Merger nor resolve hiatus, as in (24a). In the final cycle CP is interpreted and *ni-* also undergoes Phonological Merger, giving ~~*ni-gi*~~[*(nigi)(da:)(gamo)(se:)*]. Note that were the tense marker not there, the person prefix would also trigger epenthesis (*nida:gamose:*). That each of the prefixes prompts the same phonological processes independently leads to the conclusion that there are three phonological cycles within the word, and that two of them are word level. Therefore, the word-level phonology, in the terms of SOT, may iterate. The particulars of the phonological rule that can account for the distinction between deletion and epenthesis in Ojibwe will be taken up in the following section.

This section has argued that phonological rule application is not governed by a limited number of strata, and that neither morphemes nor particular morphosyntactic domains can be linked with cycle-specific phonological processes. This is predicted within a computational system where notions like stem or word are not primes in any module. It is therefore arguably simpler here to do away with the notion of strata and to remain with the notion that cyclic Spell-Out is triggered at certain points in the syntactic derivation. Any morpheme situated within a Spell-Out domain will undergo interpretation, leading to the expectation that we may find morphemes that display varying phonological behaviour dependent on their syntactic configuration in a particular derivation, an expectation that is borne out.

8.5 Future directions

In this chapter we have focused on two notions crucial to any theory of the phonology–morphosyntax interface: modularity and derivation. We have examined two very closely related theories and seen that the distinctions between them are quite nuanced. We have not delved too deeply, however, into the exact form of rules in a RBP that will give us the kinds of outputs that we see in this chapter and cross-linguistically. A theory like (S)OT has had as one of its foci an investigation of how constraints may re-rank across strata and across languages. This is due to the fundamental premise of OT that all constraints are potentially active in the grammar of each language. Rule-based theories, however, as they are not based on the contention that all rules are present in all languages, have been based upon the premise that rules may or may not be active in a particular grammar, or in a particular cycle. Any consistent cyclic effects on the form or output of rules, or patterns in the output forms at different cycles, are not derivable from the basic tenets of either OT constraint ranking or rule construction in and of themselves. Each of these theories must therefore work at motivating the patterns we do see. One evident pattern is something like the Strict Cycle Condition (41), which is akin to the Phase-theoretic notion of Phase Impenetrability (42).

(41) "<an> . . . association created in the inner domain cannot be undone in an external domain" *(Kaye 1995: 307)*.

(42) "[the phonological component] is greatly simplified if it can 'forget about' what has been transferred to it at earlier phases; otherwise, the advantages of cyclic computation are lost" *(Chomsky 2004: 107)*.

Both of these formulations are stipulative or descriptive rather than explanatory. The underlying explanation for why it should be the case that cyclic outputs should persist is therefore a domain that needs further investigation. We have seen herein that both Itô & Mester (1995) and Bermúdez-Otero & Trousdale (2012) have proposed that the different rankings of Faithfulness constraints in different parts of the grammar can explain cyclic phonological persistence. These theories take the motivation for these ranking distinctions to be controlled at least partially by extra-phonological elements like restrictions on diachronic change or the existence of co-phonologies. Here I have offered a purely phonology-internal motivation for persistence: structure-building (see Newell 2014, 2015 for further details). Phonological interpretation alters the target of a rule, but will not always eradicate the structural environment for the application of a rule, leading to divergences in rule application (see also Honeybone 2005 for a discussion of how the amount of melodic structure in the representation of a segment affects phonological rule application).

Take the hiatus resolution strategies in Ojibwe: deletion and epenthesis. At the first cycle of interpretation segments may be underspecified, and will not yet have undergone syllabification, nor will they have been organized into feet. Here a ban on hiatus is resolved through a deletion rule (remember that constraints, or bans, on certain structures are not disallowed in an RBP framework). Breaking down the timeline of, say, the construction of syllables, we are confronted with a point in the derivation where it must be determined which segments in a string have the properties that allow them to form licit nuclei (ex. sonority). Therefore, before projection of any syllabic structure, the derivation is cognizant of whether a string contains a VV sequence. At this point, deletion can occur before syllabification, and consequently no suprasegmental information is destroyed by this operation (VV→V). In the case of hiatus resolution across cycles, after Phonological Merger, the sequence of vowels is one where each of the two segments is enveloped in suprasegmental information structure (they have been syllabified). Note that Phonological Merger in Ojibwe only occurs if a prefix cannot project a licit foot. This inability can only be determined after projection of structure at the syllabic level. The sequence (V)σ(V)σ is crucially structurally distinct from the sequence VV, but similar enough in that the constraint against hiatus is still triggered. In order to conserve the previously built syllabic structure and satisfy *Hiatus, epenthesis is effected. Now, this may appear to be very close to an OT account, where constraints are pitted one against the other. It is contended here, however, that it is not the structure of a rule or a constraint that crucially distinguishes CBPs from RBPs. It is rather that the motivation for the application or non-application of operations in an RBP account must be purely phonological (features, structures), where the overt application of an OT constraint is due to extra-phonological considerations (ranking). As both RBP and CBP frameworks refer to the same phonological primes (features, syllables, etc.), then constraint ranking bears the burden of requiring independent justification. If accounts like the ones herein, where cyclic interpretation paired with structural underspecification can account for the cross-cyclic patterns attested, then this weakens the support for an explanation that calls on constraint re-ranking. Teasing apart the nuances of these types of accounts is a fertile area of investigation.

Other future research directions pertinent to the questions raised in this chapter are the following:

(1) Do we need the Prosodic Hierarchy/representational domain delimitation? Does it perform phonological duties that cannot be subsumed by cyclic, procedural derivation? Relatedly, can SOT function in the absence of alignment constraints?
(2) What are the restrictions on the destruction/modification of previously computed phonological structure? It is uncontroversial that deletion and feature-changing operations exist. This being the case, what constrains their application?
(3) Is the distinction between rules and constraints the issue here, or are the crucial distinctions between these frameworks in the different assumptions regarding the

organization of the grammar? How fundamental are notions of structure-building vs constraints-on-structure, and what are the distinctions between them?

(4) What are the pertinent cycles that determine the timing of phonological interpretation, and how are they defined?

This is obviously not an exhaustive list of future research directions related to the RBP/CBP debate at the interfaces. These questions can however help guide us to a deeper understanding of the organization of the entire grammatical computational system. Their answers can lead us to a better understanding not only of the phonological module, but also of the syntactic module that underlies it, and of the translation operations between the two. We will finish here with a short list of the advantages of an RBP theory presented in this chapter. First, eliminating reference to the Prosodic Hierarchy is desirable on theoretical grounds. The PH is not fundamental to the RBP toolbox in the same way that Align is claimed to be in CBP. Secondly, phonological cycles are not restricted in number in an RBP. We have seen that once we no longer take recourse to a Prosodic Hierarchy that this proliferation of cycles is unavoidable and, crucially, necessary. The different phonological tendencies at each cycle then need to be governed solely by phonological features and structures. Finally, cycles in an RBP don't make reference to elements that are undefinable in the syntactic component (stem, word, and even phrase in some frameworks). Syntactic and phonological cycles are determined based on identical elements: those which define a syntactic cycle, or Phase.

8.6 Further reading

Nespor, M. & I. Vogel (1986/2007) ***Prosodic Phonology/Prosodic Phonology: with a New Foreword*** This classic work on the Prosodic Hierarchy offers a detailed discussion of the motivations for the theory, as well as a trove of data relevant to the study of the morpho-syntax–phonology interface.

Selkirk, E. (2011) ***The Syntax–Phonology Interface*** This is one of the more recent updates of the Theory of the Prosodic Hierarchy (by the original proponent of the theory) that demonstrates how the view of structure–phonology relations has evolved.

Scheer, T. (2010) ***A Guide to Morphosyntax–Phonology Interface Theories: How Extra-Phonological Information Is Treated in Phonology since Trubetzkoy's Grenzsignale*** This book covers the theoretical path that theories of the interface in generative linguistics have taken in minute detail, and offers an in-depth analysis of the theoretical premises and implications of each advance.

Bermúdez-Otero (to appear) ***Stratal Optimality Theory*** Currently posted on his website as a library of separate papers, this book will cover the theory of Stratal Optimality Theory, offering a detailed look at the data that support a cyclic account of the morphosyntax–phonology interface.

Newell, H., M. Noonan, G. Piggott & L. Travis (2017) ***The Structure of Words at the Interfaces*** This is a collection of works, many of which focus on interactions at the morphosyntax–phonology interface with an eye to determining how the elements that we pre-theoretically describe as 'words' emerge at the PF interface.

8.7 Related topics

Purnell's chapter and Vaux and Myler's chapter in the RBP section, Ramsammy's chapter on the interfaces in the OT section, Charles Reiss' chapter on Substance Free Phonology, and Bermúdez-Otero's chapter on Stratal Phonology.

Notes

1 Scheer (2013) notes that this non-communication between the syntax and the phonology may be restricted to melodic primes.
2 For a more detailed account of the phonological motivation for Phonological Merger in the case of English affixes see Newell (2016a).
3 The adjectival and verbal semantics of *un-* seem to be different, but this is arguably due to the semantics of verbs and adjectives, rather than the semantics of *un-*. *un-* implies reversal: of direction of timeflow in the case of verbs, and of positive/negative scale in the case of adjectives. In the case of nouns *un-* indicates the opposite (reversal) of reference (your unbirthday is a day that is not your birthday).
4 That the stem-level affiliation of *-al* overrides the word-level affiliation of *-ment* is a separate issue that we will ignore here (but see Newell 2016b). This ability of an outer affix to trigger the reassignment of stress on an interior word-level affix is problematic for CBP and RBP accounts.

References

Adger, D. (2006) Stress and phasal syntax, *ms*. Queen Mary College, University of London.
Aissen, J. (1999) Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory* 17(4). pp. 673–711.
Avery, P. and K. Rice (1989) Segment structure and coronal underspecification. *Phonology* 6(2). pp. 179–200.
Bachrach, A. and M. Wagner (2007) Syntactically driven cyclicity vs. output-output correspondence: The case of adjunction in diminutive morphology. *U. Penn Working Papers in Linguistics* 10(1). pp. 1–16.
Baković, Eric. (2007) A revised typology of opaque generalisations. *Phonology* 24(2). pp. 217–259.
Bale, A., M. Papillon and C. Reiss (2014) Targeting underspecified segments: A formal analysis of feature-changing and feature-filling rules. *Lingua* 148. pp. 240–253.
Bermúdez-Otero, R. (2007) Marked phonemes vs marked allophones: Segment evaluation in Stratal OT. Paper presented at the Workshop on Segment Inventories, GLOW XXX, Tromsø, Apr. 11.
Bermúdez-Otero, R. (2012) The architecture of grammar and the division of labour in exponence. In J. Trommer (ed.) *The morphology and phonology of exponence* (Vol. 41). pp. 8–83. Oxford: Oxford University Press.
Bermúdez-Otero, R. (2013) *The stem-level syndrome*. UPenn Linguistics Department Speaker Series. Philadelphia 11.
Bermúdez-Otero, R. (2014) Stratal Optimality Theory. In *Oxford studies in theoretical linguistics*. Oxford: Oxford University Press. Partly available at [http://myweb.tiscali.co.uk/bermudez/Stratal_Optimality_Theory.htm].
Bermúdez-Otero, R. (2015) Paradigmatic effects without output-output correspondence: The role of lexical acquisition. Paper presented at Spell-out and the Syntax–Phonology Interface. Tromsø. Apr. 24.
Bermúdez-Otero, R. and A. Luis (2009) Cyclic domains and prosodic spans in the phonology of European Portuguese functional morphs. Presented at the 6th Old World Conference in Phonology (OCP). Edinburgh. Jan. 24.
Bermúdez-Otero, R. and G. Trousdale (2012) Cycles and continua: On unidirectionality and gradualness in language change. In T. Nevalainen and E. Closs Traugott (eds.) *The Oxford handbook of the history of English*. pp. 691–720. New York: Oxford University Press.
Boeckx, C. and K. Grohmann (2007) Remark: Putting phases in perspective. *Syntax* 10(2). pp. 204–222.
Boersma, P. (1997) How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences* 21. pp. 43–58. University of Amsterdam.
Booij, G. (1997) Non-derivational phonology meets Lexical Phonology. In I. Roca (ed.) *Derivations and constraints in phonology*. pp. 261–288. Oxford: Clarendon Press.
Booij, G. and J. Rubach (1984) Morphological and prosodic domains in Lexical Phonology. *Phonology* 1. pp. 1–27.

Borer, H. (2013) *Taking form: Structuring sense*, Vol. 3. Oxford: Oxford University Press.
Bošković, Ž. (2014) Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). pp. 27–89.
Cegalla, D.P. (2008) *Novíssima gramática da língua portuguesa*, 48th ed. São Paulo: Companhia Editora Nacional.
Chomsky, N. (1999) Derivation by Phase. MIT Occasional Papers in Linguistics, 18. MIT Working Papers in Linguistics. Department of Linguistics and Philosophy, Cambridge, Mass.
Chomsky, N. (2001a) Derivation by phase. In M. Kenstowicz (ed.) *Ken Hale, a life in language*. pp. 1–52. MIT Press.
Chomsky, N. (2001b) Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20. pp. 1–28.
Chomsky, N. (2004) Beyond explanatory adequacy. In A. Belletti (ed.) *Structures and Beyond: The Cartography of Syntactic Structure* (3). pp. 104–131. Oxford: Oxford University Press.
Chomsky, N. and M. Halle. (1968) *The sound pattern of English*. New York: Harper & Row.
Clements, G.N. (1985) The geometry of phonological features. *Phonology* 2(1). pp. 225–252.
Embick, D. (2010) *Localism versus globalism in morphology and phonology*. Vol. 60. Cambridge, MA: MIT Press.
Embick, D. (2014) Phase cycles, φ-cycles, and phonological (in) activity. In S. Bendjaballah, N. Faust, M. Lahrouchi, and N. Lampitelli (eds.) *The form of structure*, *the structure of forms: Essays in honor of Jean Lowenstamm*. pp. 271–286. Amsterdam: John Benjamins.
Embick, D. and M. Halle. (2005) On the status of stems in morphological theory. In T. Geerts, I. van Ginneken, and H. Jacobs (eds.) *Romance languages and linguistic theory 2003: Selected papers from 'Going Romance' 2003, Nijmegen*. pp. 37–62. Amsterdam: John Benjamins.
Epstein, S. and T.D. Seely (2002) Rule applications as cycles in a level-free syntax. In S. Epstein and T. Daniel Seely (eds.) *Derivation and explanation in the minimalist program*. pp. 65–89. Oxford: Blackwell.
Fabb, N. (1988) English suffixation is constrained only by selectional restrictions. *Natural Language & Linguistic Theory* 6(4). pp. 527–539.
Halle, M. and A. Marantz (1993) Distributed morphology and the pieces of inflection. In K. Hale and S.J. Keyser (eds.) *The view from building 20*. pp. 111–176. Cambridge, MA: MIT Press.
Halle, M. and A. Marantz (1994) Some key features of distributed morphology. *MIT Working Papers in Linguistics* 21. pp. 275–288.
Hayes, B. (2000) Gradient well-formedness in Optimality Theory. In J. Dekkers, F. van der Leeuw, and J. van de Weijer (eds.) *Optimality Theory: Phonology, syntax, and acquisition*. pp. 120–157. Oxford: Oxford University Press.
Honeybone, P. (2005) Sharing makes us stronger: Process inhibition and segmental structure. In P. Carr, J. Durand, and C.J. Ewen (eds.) *Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson*. pp. 167–192. Amsterdam: John Benjamins Publishing Company.
Itô, J. and A. Mester (1995) The core-periphery structure of the lexicon and constraints on reranking. In J. Beckman, S. Urbanczyk, and L. Walsh (eds.) *Papers in Optimality Theory*. pp. 181–209. Amherst: GLSA.
Itô, J. and A. Mester (2013) Prosodic subcategories in Japanese. *Lingua* 124. pp. 20–40.
Jäger, G. (1999) Optimal syntax and optimal semantics. Paper presented at DIP-Colloquium.
Julien, M. (2002) *Syntactic heads and word formation*. Oxford: Oxford University Press.
Kaisse, E.M., and P.A. Shaw (1985) On the theory of Lexical Phonology. *Phonology* 2(1). pp. 1–30.
Kaye, J. (1995) Derivations and interfaces. In J. Durand and F. Katamba (eds.) *Frontiers of phonology*. pp. 289–332. London & New York: Longman.
Kean, M.-L. (1974) The strict cycle in phonology. *Linguistic Inquiry* 5(2). pp. 179–203.
Kilborne-Ceron, O., H. Newell, M. Noonan and L. Travis (2016) Phase domains at PF: Root suppletion and its implications. In H. Harley and D. Siddiqi (eds.) *Morphological metatheory*. pp. 121–161. Amsterdam: John Benjamins.

Kiparsky, P. (1982) From cyclic phonology to lexical phonology. In H. van der Hulst and N. Smith (eds.) *The structure of phonological representations 1*. pp. 131–175. Dordrecht: Foris.
Kiparsky, P. (2000) Opacity and cyclicity. *The Linguistic Review* 17. pp. 351–367.
Kiparsky, Paul. 2007. Description and explanation: English revisited. Paper presented at the 81st Annual Meeting of the Linguistic Society of America. Anaheim. Jan. 5. Slides available at [www.stanford.edu/~kiparsky/Papers/lsa2007.1.pdf].
Lebeaux, D. (1988) Language acquisition and the form of the grammar. Ph.D. dissertation. University of Massachusetts, Amherst.
Leben, W. (1973) Suprasegmental phonology. Ph.D. dissertation. MIT, Cambridge, MA.
Legendre, G., W. Raymond and P. Smolensky (1993) An Optimality-theoretic Typology of Case and Grammatical Voice Systems. In *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society*. pp. 464–478. Berkeley.
Leu, T. (2015) Is there no homophony? Paper presented at the Meeting of the Canadian Linguistic Society. University of Ottawa. May 30.
Lowenstamm, J. (2010) Derivational affixes as roots (phasal spellout meets English stress shift), *ms*. Université Paris-Diderot and CNRS.
Marantz, Alec. (1997) No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). pp. 201–225.
Marantz, Alec. (2001) Words and things. *Ms*. MIT.
Marvin, T. (2002) Topics in the stress and syntax of words. Doctoral dissertation. MIT.
McCarthy, J. (1988) Feature geometry and dependency: A review. *Phonetica* 45. pp. 84–108.
McCarthy, J. (2000) Harmonic serialism and parallelism. *Linguistics Department Faculty Publication Series. Paper 98*.
McCarthy, J. (2003) OT constraints are categorical. *Phonology* 20. pp. 75–138.
McCarthy, J. (2007) *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox.
McCarthy, J. and A. Prince (1993a) Generalized alignment. In G. Booij and J. van Marle (eds.) *Yearbook of morphology*. pp. 79–153. Dordrecht: Kluwer.
McCarthy, J. and A. Prince (1993b) Prosodic morphology: Constraint interaction and satisfaction. *Ms*. University of Massachusetts, Amherst and Rutgers University.
Michaels, J. (2009). To alternate or not to alternate: What is the boundary? In S. Kahn, C. Moore-Cantwell, and R. Staubs (eds.) *Proceedings of North East linguistic society*. Vol. 40. pp. 93–106.
Mohanan, K. (1986) *The theory of lexical phonology*. Dordrecht: D. Reidel Publishing Company.
Myler, N. (2015) Stem storage? Not proven: A reply to Bermúdez-Otero 2013. *Linguistic Inquiry* 46(1). pp. 173–186.
Nespor, M. and I. Vogel. (1986) *Prosodic phonology*. Dordrecht: Foris.
Nespor, M. and I. Vogel. (2007) *Prosodic phonology: With a new foreword*. Vol. 28. Berlin: Walter de Gruyter.
Newell, H. (2005a) Bracketing paradoxes and particle verbs: A late adjunction analysis. In S. Blaho, L. Vicente, and E. Schoorlemmer (eds.) *Proceedings of console XIII*. pp. 249–272.
Newell, H. (2005b) A late adjunction solution to bracketing paradoxes. *Proceedings of the 35th Annual Meeting of the North East Linguistic Society (NELS)*. No.2.
Newell, H. (2008) Aspects of the morphology and phonology of phases. Doctoral dissertation. McGill University.
Newell, H. (2014) Phonological persistence. Paper presented at the North American Phonology Conference (NAPhC). Concordia University. May 9–11.
Newell, H. (2015) Structural sensitivity in phonology: Phonological persistence. Paper presented at the 12th Old World Conference in Phonology (OCP). Barcelona. Jan. 30.
Newell, H. (2016a) The ontology of English morpho-phonology. Paper presented at the 24th Manchester Phonology Meeting (MfM). University of Manchester. 26–28 May.
Newell, H. (2016b) English Lexical Levels are not Lexical. Ms. UQAM.
Newell, H., M. Noonan, G. Piggott and L. Travis (eds.) (2017) *The structure of words at the interfaces*. Oxford: Oxford University Press.

Newell, H. and G. Piggott (2014) Interactions at the syntax–phonology interface: Evidence from Ojibwe. *Lingua* 150. pp. 332–362.

Newell, H. and T. Scheer (2007) Procedural first. Paper given at the 38th Poznań Linguistic Meeting, Gniezno, Sept. 13–16. Available at [www.unice.fr/dsl/tobias.htm].

Nissenbaum, J. (2000) Investigations of covert phrase movement. Doctoral dissertation. MIT.

Ochi, M. (1999) Multiple spell-out and PF adjacency. *Proceedings of the North East Linguistic Society* 29.

Pater, J. (2000) Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17(2). pp. 237–274.

Pesetsky, D. (1985) Morphology and logical form. *Linguistic Inquiry* 16. pp. 193–246.

Prince, A. and P. Smolensky (1993) Optimality Theory: Constraint interaction in Generative Grammar. *ms*.

Prince, A. and P. Smolensky (2004) *Optimality Theory: Constraint interaction in generative grammar*. Hoboken, NJ: John Wiley & Sons.

Sagey, E. (1986) The representation of features and relations in non-linear phonology. Doctoral dissertation. MIT.

Scheer, T. (2004) *A lateral theory of phonology*. Vol. 1. Berlin: Mouton de Gruyter.

Scheer, T. (2008) Why the prosodic hierarchy is a diacritic and why the interface must be direct. In J. Hartmann, V. Hegedűs, and H. van Riemsdijk (eds.) *Sounds of silence: Empty elements in syntax and phonology*. Vol. 63. pp. 145–192. Leiden: Brill.

Scheer, T. (2010) *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin: Walter de Gruyter.

Scheer, T. (2012) Chunk definition in phonology: Prosodic constituency vs. phase structure. In *Modules and Interfaces*. pp. 221–253. Lublin: Wydawnictwo KUL.

Scheer, T. (2013) Melody-free syntax. Paper presented at the Workshop on complexity at UQAM, Montréal. Feb. 8.

Selkirk, E. (1981) On the nature of phonological representation. *Advances in Psychology* 7. pp. 379–388.

Selkirk, E. (1984) *Phonology and syntax*. Cambridge, MA: MIT Press.

Selkirk, E. (1996) The prosodic structure of function words. In J.L. Morgan and K. DeMuth (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. pp. 187–214. Mahwah, NJ: Lawrence Erlbaum Ass. Inc.

Selkirk, E. (2011) The syntax-phonology interface. In J.A. Goldsmith, J. Riggle, and C.L. Alan (eds.) *The handbook of phonological theory*. Vol. 75. Hoboken, NJ: John Wiley & Sons.

Starke, M. (2010) Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). pp. 1–6.

Stepanov, A. (2001) Late adjunction and minimalist phrase structure. *Syntax* 4(2). pp. 94–125.

Svenonius, P. (2004) On the edge. In D. Adger, C. de Cat, and G. Tsoulas (eds.) *Peripheries: Syntactic edges and their effects*. pp. 259–287. Dordrecht: Kluwer.

Travis, L. (2010) *Inner aspect: The articulation of VP*. Vol. 80. Berlin: Springer Science & Business Media.

Truckenbrodt, H. (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30. pp. 219–255.

Uffmann, C. (2011) The organization of features. In M. van Oostendorp, C.J. Ewen, E. Hume, and K. Rice (eds.) *The Blackwell companion to phonology*. Blackwell Publishing. Blackwell Reference Online. [www.blackwellreference.com/subscriber/tocnode.html?id=g9781405184236_chunk_g978140518423629].

Uriagereka, J. (1999) Multiple spell-out. In S. Epstein and N. Hornstein (eds.) *Working minimalism*. pp. 251–282. Cambridge, MA: MIT Press.

Vaux, B. (2008) Why the phonological component must be serial and rule-based. In B. Vaux and A. Nevins (eds.) *Rules, constraints, and phonological phenomena*. pp. 20–60. Oxford: Oxford University Press.

Wiltschko, M. and R.-M. Déchaine (2010) Interface Syntax, *ms*. University of British Columbia.

Zwicky, A.M. and G.K. Pullum (1986) The principle of phonology-free syntax: Introductory remarks. *The Ohio State University Working Papers in Linguistics* 32. pp. 63–91.

9
Government Phonology
Element Theory, conceptual issues and introduction

Tobias Scheer and Nancy C. Kula

9.1 Introduction and conceptual issues

9.1.1 What you get is not what you see

Government Phonology (GP) grew out of developments of autosegmental phonology that characterized phonological research in the 1980s. Work by Jonathan Kaye and Jean Lowenstamm during that period (e.g. Kaye & Lowenstamm 1981, 1984) was condensed into what today is called Standard Government Phonology. A first step in 1985 concerned the internal structure of segments (Kaye et al. 1985; see sections 9.1.4.3 and 9.2.1, below) and a second step in 1990 regarded syllable structure (*Phonology Yearbook* 7.2 guest-edited by Jonathan Kaye, containing namely Kaye et al. 1990). There are three book-length presentations of the theory: Charette (1991); Harris (1994); and Gussmann (2002). The framing of GP within the architecture of grammar and the larger landscape of Cognitive Science is described in Kaye (1989), *Phonology: A Cognitive View*. Finally, there are two short guides to standard GP written from hindsight, by Kaye (2000, unpublished but available online) and Scheer (2004: §623).

GP as such may be divided into several periods. In an initial period, what today is called standard GP defined the conceptual essentials and accordingly built first versions of sub-segmental and syllable structure. This period closed with the publication of Charette (1991), which provides a snapshot of the state in which GP was then. This coincided more or less with the end of the Montreal group where Jonathan Kaye and Jean Lowenstamm had collaborated: Kaye moved to London, Lowenstamm to Paris. The former further developed sub-segmental aspects, the latter syllable structure. The Revised Theory of Elements then emerged from the work of Kaye and Charette at the School of Oriental and African Studies (SOAS) in London (whereby GP 2.0 is a more recent offspring thereof, following the same idea that the number of melodic primes need to be reduced), while Lowenstamm introduced CVCV (or Strict CV).

The choice of the word *cognitive* in the title of Kaye (1989) is programmatic: it was meant to take issue with so-called concrete approaches (Tranel 1981) which at the time opposed SPE-type abstractness, and with anything that grounds phonology in phonetics (a chapter in the book is called "The Nonphonetic Basis of Phonological Phenomena"). For Kaye, phonology is a computational system (i.e. operating an input-output transformation based on a stored set of instructions) that, much in the spirit of Saussure's Langue, is self-contained and operates on purely cognitive units. These must relate to phonetic realizations, but the relationship is not one-to-one and the phonological (cognitive) identity of a unit cannot be predicted or derived from its phonetic properties. Even though phonological and phonetic properties sometimes coincide, the analyst will often be fooled when trusting phonetics: in French for example word-initial [w] may allow elision (*l'ouate* [lwat] where the definite article /lə/ is followed by "cotton wool") or may not (*le watt* [lə wat] where the definite article /lə/ precedes "watt"), despite the fact that the two words are phonetically identical ([wat]). Yet the two [w]s must be distinct. Kaye & Lowenstamm (1984) suggest that this difference is syllabic, rather than melodic: [w] belongs to a complex nucleus in *ouate* (and therefore a hiatus leading to elision is created), but to an Onset in *watt* (no hiatus occurs since the two vowels are separated by a consonant, the [w]).

This is why caution demands that phonological units never be based on their phonetic properties. Rather, the only source of phonological knowledge is phonological *behaviour* (this is what Kaye 2005: 283 calls the phonological epistemological principle). We have seen that the syllabic identity of French word-initial [w] is only betrayed by its behaviour (regarding elision). In the same way, the fact that [ɛ] is phonetically front does not entitle the analyst to deduce that it contains a phonological prime encoding frontness (e.g. [-back] in binary systems, |I| in unary approaches). But the fact that an [ɛ] causes palatalization of a preceding velar consonant (and to the extent that this process is managed by phonological computation, on which see section 9.1.4) does allow us to conclude that this vowel is phonologically front. Palatalization is a process denoting phonological behaviour, [ɛ] is a static pronunciation.

In practice, of course, this does not prevent the analyst from proceeding by trial and error, i.e. from making hypotheses based on static phonetic properties that turn out to be wrong. In Polish (and many other Slavic languages), for example, some [ɛ]s palatalize, others do not: given the nominative sg *lo*[t] "flight", compare *lo*[t͡ɕ]-[ɛ] "id., locative sg" with *lo*[t]-[ɛ]*m* "id., instrumental sg". Phonological behaviour thus tells us that the palatalizing locative [ɛ] must possess a frontness prime – but it does not provide any clues to the phonological identity of the non-palatalizing instrumental [ɛ]. Concluding that the frontness prime is absent from the phonological makeup of this [ɛ] is one option, but depending on additional evidence from the language, other analyses may be entertained: in Element Theory terms, Gussmann (2007: 56ff.) for example proposes that in Polish palatalizing [ɛ] identifies as I̲-A (Element Theory is introduced in section 9.2.1, below: |I| is the melodic prime representing frontness, |A| denotes the low position of the tongue, heads are underscored), while non-palatalizing [ɛ] realizes the empty-headed expression _-I-A (where "_" represents the empty head). Palatalization, then, is sensitive not to the presence of I, but to its status as a head or non-head (this distinction is introduced in greater detail in section 9.2.1, below).

9.1.2 Modular architecture

9.1.2.1 Spell-out and phonetic interpretation

Let us continue to examine the Polish pattern. What about the other analytical option whereby the frontness prime is absent from the phonological expression of a non-palatalizing front vowel? In such a scenario, why would this kind of vowel be pronounced as front? Is there a source of phonetic frontness other than a lexical or a phonological specification? The answer is yes, and this has to do with the way GP conceives the interface of phonology with phonetics (for more detail see Chapter 11.2). Since *Aspects* (Chomsky 1965: 15ff.), the generative architecture of grammar in form of the inverted T model (which is reproduced on the first page of Kaye et al. 1990, henceforth KLV 1990) is modular in the Fodorian sense (Fodor 1983) and consists of three independent computational systems.

In production, first morpho-syntax concatenates lexical items retrieved from long-term memory, then phonology and semantics interpret the result of this concatenation process. Since different computational systems operate on distinct vocabulary and are unable to parse ("understand") the idiom of their neighbours (a property called *domain specificity* in Cognitive Science, e.g. Segal 1996), they can only communicate through a translational process. Therefore, when morpho-syntactic structure is to be interpreted by phonology, it is translated into phonological vocabulary by so-called vocabulary (or lexical) insertion (e.g. Embick & Noyer 2007). This is undisputed in generative linguistics: a spell-out operation transforms portions of the morpho-syntactic tree (featuring categories such as number, person, animacy etc.) into phonologically meaningful units such as labial, occlusion etc. through a lexical (and thus arbitrary) specification (e.g. in English, past tense ↔ *-ed*).

GP takes the modular architecture of grammar seriously and holds that the interface with phonetics works exactly in the same way: a spell-out operation assigns a phonetic value to phonological primes through a lexical (and thus arbitrary) specification. This is called *phonetic interpretation* (Harris & Lindsey 1990, 1995: 46ff., see Chapter 11.2.1). In our Polish example, two lexically stored specifications assign a pronunciation to the two [ɛ]s: I̲-A ↔ [ɛ] and _-I-A ↔ [ɛ]. That is, the language pronounces two phonologically distinct items alike. This hard-wired and language-specific knowledge must be acquired by the child and has the same status as other parametric specifications, for example defining sound inventories.

Now let us come back to our original question: if a non-palatalizing front vowel lacks the phonological prime for frontness, why is it pronounced front? The answer is that spell-out may decide so. Assume a non-palatalizing [i] which phonologically identifies as an empty-headed structure without additional primes (_). Cast in surface terms (for the sake of exposition), this [i] is in fact /ɨ/ and therefore does not palatalize. It is pronounced [i], however, because there is a spell-out instruction specifying _ ↔ [i]. Dresher & Compton (2011: 222) describe Inuit dialects that have a palatalizing and a non-palatalizing [i] whereby the former (so-called strong i) comes from Proto-Eskimo i, while the latter (weak i) is a reflex of Proto-Eskimo schwa. They show that weak i not only does not palatalize, but is also subject to other processes such as assimilation, dissimilation and deletion. Dresher & Compton therefore conclude that weak i has no phonological substance: it is synchronically empty just like it was in Proto-Eskimo. The diachronic change, then, is only in the spell-out relation: the modern ([i]) and the ancient ([ə]) pronunciation realize the same phonological unit, schwa. That is, the diachronic evolution concerned the spell-out of that unit (_ ↔ ə > _ ↔ i). Its phonology remained untouched and is exactly the same in the modern and the older variety of the language. That is, there was no evolution that would have changed the phonological identity of the vowel in question: _ (↔ ə) > I (↔i) did not occur.[1]

The take-home message of GP, then, is that phonology does not work along the statement *what you get is what you see*: phonetic properties of a sound do not allow the analyst to conclude on either its syllabic affiliation (French elision) or its melodic identity (Polish/Inuit palatalization). Only their behaviour does.

9.1.2.2 Structural Analogy

The modular setup does not mean that different computational systems do not share any properties. GP is known for applying syntactic mechanisms to phonology, thereby joining the tradition of Structural Analogy developed in Dependency Phonology by John Anderson (1985, 1992 and following). In the formulation of van der Hulst (2000: 209), “grammar recapitulates, rather than proliferates, structures and principles”. This takes on the following form in GP.

(1) “What is at stake here goes well beyond a mere search for interesting or suggestive similarities. Rather, if (some of) the same principles can be shown to underlie phonological as well as syntactic organisation, the idea that such principles truly express special, idiosyncratic properties of the mind (such as the kind of asymmetries typical of natural language) will be correspondingly strengthened.”

KLV (1990: 194)

Accordingly, GP has imported several mechanisms from syntax, some of which are introduced below: the Empty Category Principle, Proper Government, c-command, a phonological version of the Minimality Condition, the Projection Principle and Structure Preservation.

9.1.3 The purview of phonology: small is beautiful

A question as old as phonological analysis is what exactly counts as phonology (or, to be precise, as phonological computation). This is what Ricardo Bermúdez-Otero (p.c.) calls the Holy Grail of phonology (Bermúdez-Otero 2007; Scheer 2015 provide overviews of the question). This issue is paramount and for obvious reasons must be decided before one can begin to build phonological theory: if you decide that your theory needs to account for a set X of empirical phenomena, it will end up wildly different according to the size of X. To anticipate, GP argues that only a very small subset (perhaps 10%) of what SPE held to be phonological is indeed managed by phonological computation. Much like current syntactic theory whose minimalist orientation also drastically reduces the set of truly syntactic phenomena, GP thus believes that *small is beautiful*.

Central to the calibration of X is the notion of overgeneration, both for GP and in the historical development of generative phonology during the so-called abstractness debate of the 1970s that was triggered by Kiparsky (1968–73). An SPE-type rule can describe any phonological process and its reverse, i.e. the set of existing processes (e.g. k → t͡s/__i,e) as much as all processes that are never produced by natural language (e.g. p → ŋ/__i). The basic ambition of the generative enterprise, as indicated by its name, is to generate *all and only* those linguistic expressions that occur. Therefore, a computational system like SPE that on top of all occurring processes can also generate all non-occurring events is in trouble. GP takes this problem very seriously at all levels: regarding the melodic makeup of segments (see section 9.2.1), syllabic constituency and the set of processes that instantiate phonological computation.

Note that there are also modern representatives of the reverse philosophy, i.e. those who argue that overgeneration is not a bad thing to have and indeed a property of all natural systems (only a small subset of what biological evolution can in principle produce is actually attested by past or present organisms). Hale & Reiss (2008) represent this SPE-defending line of thought as *big is beautiful*.

In practice, an alternation of some segment in two instances of the same morpheme may in principle represent one of the following situations.

(2) Alternations are produced by
  a. distinct lexical entries
  b. morpho-phonology
  c. allomorphy
  d. analogy
  e. phonology

For the sake of illustration, let us consider English velar softening, whereby the velars [k], [g] seem to be turned into [s], [d͡ʒ] before [i] (*electri*[k] – *electri*[s]-*ity*, *analo*[g]*ue* – *analo*[d͡ʒ]-*y*). The analytic option under (2a) considers both *electric* and *electricity* as single, i.e. morphologically non-complex, lexical entries whose pronunciation requires no concatenation and no phonological (or other) computation at all. The second analytical option (2b) refers to a computational system distinct from phonology that is entertained in traditional grammar as well as in structuralist and early generative approaches: morpho-phonology. Here computation takes into account morphological as much as phonological information (see Gussmann 2007: 10ff. for an overview, also historically). Velar softening is morphologically conditioned: it does not apply morpheme-internally (*king* is [k]*ing*, not [s]*ing*) and goes into effect only before a subset of i-initial suffixes, which Kiparsky (1982: 40f.) identifies as class 1 suffixes: -*y*, -*ity* and -*ism* (compare with class 2 suffixes such as -*ing* in *hik-ing*, which is not **hi*[s]-*ing*). Therefore velar softening qualifies for being managed by morpho-phonological computation (2b).

A third option is allomorphy (2c): there are two lexical entries *electri*[k]- and *electri*[s]- which are selected (through morphological computation) by purely morphological information (class 1 suffixes choose the latter, class 2 suffixes the former). Analogy (2d) also performs computation, but of a kind that some believe lies outside of grammar since it requires comparison with unrelated lexical items: *electri*[s]-*ity* has an [s] because the speaker knows that there are a number of other words that end in -*sity* and therefore modifies the lexical entry *electri[k]-* to become *electri[s]-*. Finally, phonological computation as under (2e) is another analytical option, whereby a grammatical instruction stored in long-term memory (a rule or constraint set) transforms [k] into [s] before [i] (in appropriate morphological contexts).

How could the analyst discover which mechanism controls particular alternations? As a consequence of the abstractness debate, much effort was put into establishing a set of formal criteria (called the "evaluation measure" or "evaluation metrics" in the 1970s, e.g. Kiparsky 1974) that allows us to decide whether an item that the analyst identifies as morphologically complex is really considered as such by the grammatical system, and if so, whether or not its computation is phonological in kind. All attempts failed, and the result in the early 1980s was a situation where everybody agreed that the position of SPE on the extreme big-is-beautiful end of the scale is not realistic. The whole thrust of the abstractness debate was to reduce this computational abstractness, shifting the analysis of a fair amount of alternations to other mechanisms. The question was to what extent the computational balloon of SPE needed to be deflated. Natural Phonology (Stampe 1972) and Natural Generative Phonology (Hooper 1976) were built around this question in the 1970s, and Lexical Phonology (Kiparsky 1982) may be argued to be an attempt to save the basic SPE architecture by shifting labour from phonological computation to the cyclic system (see Scheer 2011).

As mentioned above, in this landscape GP takes a radical position on the extreme small-is-beautiful end of the scale (akin to the natural phonologies; see Scheer 2015): most of what SPE held to be phonological are instances of some other mechanism among those mentioned under (2). In practice, then, what are the criteria that allow the analyst to decide whether a given alternation is the result of phonological computation or not? Standard GP decides along the lines under (3):[2] if an alternation is characterized by any one of the three properties, it lies outside of phonology.

(3) An alternation cannot be phonological
  a. if it is not 100% regular, i.e. surface-true OR
  b. if it has conditioning factors that are morphological and cannot be expressed by domain structure OR
  c. if there is no plausible causal relationship between the change observed and the triggering context

The proviso under (3a) is a consequence of Kaye's (1992: 141, 1995: 291) take on computation (on which more in section 9.1.4 below): phonological processes apply whenever the conditions that trigger them are met. This statement needs to be complemented with the fact that in Kaye's view there is only one set of phonological instructions. Lexical Phonology (Kiparsky 1982) has introduced the idea that different chunks of the linear string (which follow inside-out embedding and interpretation), called cycles, may be subject to different sets of rules, i.e. different phonologies. Hence given a (cyclic) structure [[[A]B]C], different computational instructions may apply to A (inner cycle, e.g. corresponding to the root), AB (intermediate cycle, e.g. corresponding to the stem) and ABC (outer cycle, e.g. postlexical phonology). There is a (depleted version of) cyclic structure in GP (the domain structure mentioned under (3b), on which more in Chapter 11.1.2), but all cycles/domains are computed by the same phonology. Hence all chunks of the linear string, wherever they occur in the cyclic/domain structure, will be subject to this unique computation,[3] which means that in the presence of a phonological k-to-s-before-i instruction that instantiates velar softening, all instances of the linear string will be modified. That is, there must not be any surface [ki] sequence at all: monomorphemes like *king* are impossible.

Velar softening cannot be phonological for yet another reason: there are exceptions such as *monar*[k] – *monar*[k]-*ism*, *patriar*[k] – *patriar*[k]-*y*. In the same way, it fails the productivity test with words like *Iraq*: native speakers seem unable to even parse *Ira*[s]*ity* ("the property of being typically like Iraq"), but are able to make sense of *Ira*[k]*ity*.

Finally, (3c) is called non-arbitrariness in GP and appears in KLV (1990: 194, see also Gussmann 2007: 30ff.; Pöchtrager 2006: 19ff.). The front vowel [i] is certainly a plausible candidate for triggering palatalization (unlike, say, [a] or [u]). But the fact that the other front vowels do not have the same effect begs the question. Also, the structural change itself needs to be inspected: [i] may be a good palatalization trigger, but a trigger for which palatalization exactly? Cross-linguistic (and also diachronic) experience leads the analyst to mistrust [k] → [s], since typical results of palatalization are [t͡s] and [t͡ʃ], or maybe [c]. At a previous diachronic stage of English, the palatalization at hand was actually [k] → [t͡s], but the (unconditioned) loss of affricates ([t͡s] > [s]) turned it into a suspicious alternation. This is the typical way for regular processes to become opaque while aging (see Bach & Harms 1972): through a context-free change that affects the input or the output of a process. The question, then, is whether phonological computation, which managed the original alternation, will be able to accommodate this kind of evolution. GP says no: it breaks down because the new alternation is no longer expressible in phonological terms.

It needs to be realized (though not always is by GP practitioners) that non-arbitrariness is theory-dependent: there is no pre-theoretical or surface-based definition of what exactly is "expressible in phonological terms". Depending on the theory of melodic representations, a front prime that enters a velar stop and turns it into an [s] may or may not be something that can be described. Given the initial melodic representations of GP, it cannot, and hence velar softening is rejected as being beyond the realm of phonology. The next question, then, is how the melodic representations of GP have come into being, and the answer of course is that they are designed in order to account for a number of phonological processes. Which means that there was a selection among all alternations, some of which were judged to be truly phonological (and hence informed melodic theory), while others were left aside. Rather than being circular, this is the regular dialectic between data (bottom-up) and hypothesis (top-down) that is the common ground of scientific activity. But this means that the criterion under (3c) does not judge alternations per se: it assesses them only given an already established phonological theory. And this theory of course can (and should) change over time, which may consider a previously non-phonological alternation phonological, or vice-versa.

The representation of velar consonants is an example of how a change in the theory has modified the set of alternations that are held to instantiate phonological computation. In the first version of consonantal representations in GP (Harris & Lindsey 1995: 67), velars do not contain the element |U|: they are empty-headed (see section 9.2.1.1). Therefore the variation of the Czech vocative (masc.) marker for example must be declared non-phonological: -*i* attaches to palatal-final stems (*Tomáš-i* "Thomas Vsg"), -*u* occurs after velar-final stems (*František-u*) and -*e* is found elsewhere (*Jakub-e*, *Milan-e*). Since u is made of |U| but velars do not contain this element, they cannot spread it to the suffix. Today most elemental approaches to the internal structure of consonants acknowledge the presence of |U| in velars (Backley 2011: 79ff.; see section 9.2.1), which means that the Czech vocative alternation is considered phonological in kind.

9.1.4 Computation in GP

9.1.4.1 Anti-serialism

Rooted in the properties of the universal Turing machine (Turing 1936–37), serialism (or derivationalism) lies at the heart of the standard theory of Cognitive Science that emerged in the 1950s, and whose application to linguistics produced generative grammar (see e.g. Gardner 1985). Serialism is the idea that computation in the mind involves a set of instructions that act on the input in such a way that it experiences step-by-step modifications which occur in a chronological and logical order where the output of step n-1 is the input to step n.

In generative grammar, serialism shows up as extrinsically ordered rules in phonology, and in early syntax as extrinsically ordered transformations. The latter were abandoned in the early 1980s when Government Binding (GB) theory introduced so-called move α, a system where movement (computation) is free in itself, but controlled by constraints on representations (e.g. Newmeyer 1986: 163ff.). Move α represents an important turn in syntactic theory away from restrictions on computation itself (Chomsky's 1973 original Strict Cycle Condition, extrinsically ordered transformations) in favour of a central role of well-formedness constraints on representations such as barriers, the ECP, case checking and so forth. The development of autosegmental structures in the 1980s follows the same track:

representations are governed by well-formedness conditions such as the OCP or no line-crossing. While generative syntax thus abandoned serialism in 1981, the representational blossoming of the early 1980s left extrinsically ordered rules untouched in phonology: well-formedness conditions were added on top of them, and the result was a hybrid model (see Scheer 2011 for more detail). In the second half of the 1980s, though, a general discomfort with serialism arose, which was pervasive through the entire field.

GP participated in the anti-serialist movement, considering that extrinsic rule ordering was empirically vacuous. That is, examples where serial ordering of instructions is alleged to be critical are either based on erroneous data, involve misanalysis or concern processes whose properties disqualify them as instances of phonological computation. An example for erroneous data is Martin Joos' famous dialect B of Canadian English for which there is no evidence (Kaye 1990, 2008), but which was used by Bromberger & Halle (1989) as the litmus test for rule ordering. Examples for processes that are not phonological in nature are Trisyllabic Shortening (or other traces of the Great Vowel Shift) and the aforementioned velar softening.

While anti-serialism is an important feature of the early identity of GP, there is almost no trace of this programme in print: Kaye (1990) is only a brief comment about the non-existence of dialect B, and the GP literature of the 1980s (Kaye et al. 1985: 305; Lowenstamm & Kaye 1986: 97) heralds the programmatic claim that there are no rules (GP is a "no-rule approach"), but essentially leaves it at that. Kaye et al. (1985: 305) merely explain that they expect Principles and Parameters theory coming from the then freshly established GB syntax to take over 100% of the function of ordered rules (just like in syntax). They add that "at the moment of writing, this view of phonology remains a long-term objective of our research programme". The following section locates this ambition in the context of the time and fleshes out the little that the further GP literature contains about how exactly computation works.

9.1.4.2 Weakly developed constraint-based computation

The latent antipathy against serialism of the late 1980s reshaped the landscape in the early 1990s by producing three theories that are based on the anti-derivational mantra: Optimality Theory, Declarative Phonology and GP. In these approaches, computation is based on constraints. Constraints, however, do not have the same status in the tree theories: while they are ranked and violable in OT, they are absolute (i.e. non-violable) in Declarative Phonology.

GP is often referred to as a representation-oriented theory of phonology, and there is certainly good reason for this characterization. A correlate of the representational focus is the fact that the programmatic statement mentioned lain aside, there is not much to be found about how computation works and what a computational instruction looks like. The only indication that was available until the mid-1990s is Kaye's (1992: 141, 1995: 291) statement according to which processes "apply whenever the conditions that trigger them are satisfied". The constraint-based character of computation in GP has appeared only since the introduction of Licensing Constraints (Charette & Göksel 1994, 1996; Kaye 2001). To date, Gussmann's (2007) book on Polish appears to be the only detailed application of this constraint-based approach.

Constraints in GP thus apply whenever a form may be modified by them, but with no extrinsic ranking or ordering (hence unlike OT), and without being violable: the set of constraints (the φ-function in Kaye's 1995 terms) is (simultaneously and) iteratively applied to the string that is submitted to interpretation, and computation ends when no further modification can be made (this is a parallel with Harmonic Serialism and OT-CC).

Expressed using serial vocabulary, this system is thus able to handle a feeding relationship (the conditions for the application of a constraint are created by the modification of the input string by another constraint), but no other (i.e. bleeding, counter-feeding, counter-bleeding). A difference must therefore be made between serial computation (GP computation is serial in the sense that constraints may apply to the same string several times, and that intermediate steps may thus exist) and serialism per se (there is no extrinsic or logical ordering of instructions, i.e. classical extrinsic rule ordering).

Also, there is no ranking or prominence (dominance) relationship among constraints: all instructions are equally important, and there is no selective application. That is, all instructions of the φ-function apply when a string is computed: there is no way for just a subset of the phonological processes to apply at a given time and to a given string. Phonology is not divisible, and "do phonology!" (the way Kaye 1995 describes the application of the φ-function) means "do *all* the phonology!".

In sum, computation has not been a central focus of GP, nor could it be said that GP has provided major contributions to computational theory. Speaking in economic terms, the behaviour of GP in the overall landscape is counter-cyclic. Writing at the peak of the representational (autosegmental) wave and having observed the see-saw movement of phonology between process- and representation-oriented extremes in the history of phonology in the 20th century, Anderson (1985) extrapolated what would come next: another round of the computational extreme, in counteraction to the representational excess. Phonology was thus programmed to produce OT, which indeed entered the scene a couple of years later. Anderson predicted its arrival, but little did he know how extreme this round of "phonology is computational and nothing else" would get. When the mainstream thus set out to shift the labour that was done by representations onto computation (constraint interaction), leaving only a decorative role for the former (de Lacy 2007 for example is explicit on the intention and reality of this movement; see Scheer 2010), GP bet on the exact reverse setup: phonological patterns are best explained by a rigid representational theory augmented with well-formedness constraints. Computation was secondary in this project, and this is reflected by the fact that little was done to explain how computation works in the GP literature.

Today things have swung back more into a midfield position, i.e. the one that Stephen Anderson (1985) argued for: a sound theory of phonology needs a theory of representations *and* a theory of computation, and these need to be independent (none must be the slave of the other). In mainstream OT, the decorative remnants of representations used to be "emerging", i.e. the result of constraint interaction (see Scheer 2010). But since the mid-2000s, there is a body of OT-based literature arguing for a return to representations that are not just a toy (or the slave) of computation: they may e.g. be hard-wired by constraining Gen (Blaho et al. 2007; Oostendorp 2005; and others).

9.1.4.3 Independence of representations and computation

While a good case can be made (and is typically made by representatives of GP) for the interdependence of pieces of representational structure (piecing together a representation with items from GP and other theories produces a monster), the same cannot be said about representations and computation. I am unaware of evidence showing that the representations of representational theory X can only be managed by the computation of computational theory A (while B does not qualify, or while Y is incompatible with A). As long as the dualistic Anderson-standard is respected (there are theories of representation and of computation, and one is not the slave of the other), the accuracy of representational and computational theories appears to be a matter of independent evaluation (Scheer 2010).

For GP, this means that the core of its contribution to phonological theory is representational in kind, and that the management of GP representations is to be sought by the best computational theory around (which may or may not be the one that is based on GP's initial take on computation). Hence there are attempts to combine GP representations with OT computation (Polgárdi 2009). Also, the small-is-beautiful positioning of GP at the extreme end of the scale describing the number of alternations that are phonological (see section 9.1.3) partly depends on how computation works: namely the decision that there is only one phonology (i.e. one set of phonological instructions) does not follow from anything. Computation could as well be chunk-specific (root, stem, postlexical) or morpheme-specific (indexed constraints, cophonologies in OT): these options substantially increase the generative power of the system and hence augment the volume of alternations that can be described. Rejecting them is a matter of an independent decision that needs to be motivated.

An important related aspect that usually goes unnoticed is the fact that the granularity of the instructions that are developed by practitioners who implement computation into explicit formal statements is much greater than what is usually done in GP. That is, in GP computation is typically referred to in prose statements such as "and then the element spreads from the nucleus to the Onset" or "and then the suffixal vowel that sits in the final empty nucleus governs the preceding nucleus, which therefore remains unpronounced (government)". OT-type constraints would simply cast these prose statements in a more formalized and more fine-grained vocabulary.

9.2 Element Theory

9.2.1 Background

It is a generally accepted position in generative phonology since Jakobson et al.'s (1952) work that segments are decomposable into smaller units. For Jakobson et al. (1952); Jakobson & Halle (1956); and subsequently Chomsky & Halle (1968), these smaller units are features with the role of capturing both the distinctive properties of segments and the various phonological operations segments are involved in. While Jakobson and Halle used primarily acoustically based features, Chomsky and Halle developed more articulatory-based features. In both systems features are treated as having a polar opposition offering a positive (+) and negative (−) value of each feature. The number of oppositions offered in each system varies between twelve (in Jakobson & Halle 1956) to over thirty (in Chomsky & Halle 1968). Within the polar opposition of features both oppositions for some features like [+Voice] can be argued to be active in grammars to account for voicing assimilation, for example, while for a number of others like [+Nasal], one opposition [+Nasal] is more active than the other [−Nasal]. In fact the latter is not seen to be active at all with no processes targeting segments that are specifically [−Nasal].

The other property that segmental sub-units such as features aim to capture is the idea of natural classes, namely that phonological processes are seen to apply to a select set of features as triggers or undergoers to the exclusion of others. Again in a feature system this works well for some select features but not at all for others. There are two main issues in classic feature theory as developed in SPE that have led to alternative perspectives on sub-segmental units: the overgeneration problem and the lack of inherent natural class predictiveness. The greater the number of features that are assumed, the larger the set of phonological segments that are predicted to exist. Within the overgeneration problem lies the central question of whether sub-segmental units should have binary oppositions or be monovalent. As a way of dispensing with these concerns GP argues that segments are composed of elements which are acoustically defined monovalent cognitive units supporting a hearer-oriented phonological grammar. This crucially distinguishes elements from features that are defined based on articulation and is in this respect more in line with Jakobson. Other comparable approaches adopting non-feature-based sub-segmental units are Dependency Phonology (Anderson & Jones 1972; Anderson & Ewen 1987) and Particle Phonology (Schane 1984).

Anderson & Jones (1972) is also significant in introducing unary features in response to some of the challenges already raised above. This was quickly adopted in non-feature-based work but also more recently in feature-based representations as, for example, surveyed in Hall (2007). The significant difference between such unary feature systems and element-based representations is the ability to be independently pronounceable, which remains a property of only the latter system. Thus in element-based representations a segment may consist of only one prime, attesting to the independent interpretability of elements, while this is never the case for features.

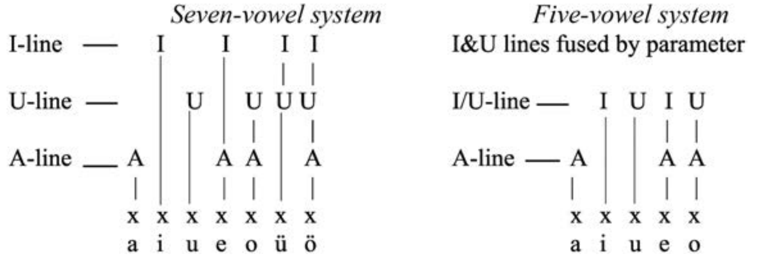
9.2.1.1 Elements and charm theory: the early days

Elements in GP were introduced in Kaye et al. (1985) where they more specifically elaborate on the representation of vowels. In this early work elements are defined as consisting of fully specified feature matrices but features are seen as only providing the phonetic interpretation of otherwise autonomous and independently pronounceable phonological units. The elements |A| |I| |U| are given as the three vowel elements from which other vowels can be derived by a combination of these basic elements. |A| |I| |U| are themselves independently interpretable as the segments /a/, /i/ and /u/. Within KLV's (1985) element representations the feature matrices of each element consisted of what they termed a *hot feature* which identified the marked feature in each matrix reflecting the dominant characteristic of each element. |A| |I| |U| had [−high], [−back] and [+round], respectively, as hot features. In this sense elements were embedded in markedness theory. In KLV (1985) elements are treated as residing on autosegmental tiers which connect with skeletal points in order to give the eventual representation of a segment. Combinatory possibilities between elements are regulated by the tiers on which they reside. Thus only elements that are on different tiers can combine while those on the same tier cannot. Tiers are defined according to the hot feature for which a particular element must be specified. In this sense the constraints that hold on individual vocalic systems are defined in the tier representations with elements that cannot combine in a particular language treated as conflating their tiers. Fused |I| and |U| tiers (Back and Round lines) are treated as the unmarked option in vowel systems based on the empirical distribution of vocalic systems. Elements on the same tier, like |I| and |U| in systems where the unmarked holds, cannot combine and OCP effects apply so that in contrast to Particle Phonology, a single element cannot be present more than once in the representation of a segment. Thus, while (4a–b) are possible compound representations, (4c) is not, with the latter represented as a single element attached to a single skeletal point. Tier conflation is depicted in (5) drawn from Cyran (1997), respectively depicting a seven- and a five-vowel system.

(4) Element combinations

a.	|A| + |I|	→	|A.I| /e/
b.	|A| + |U|	→	|A.U| /o/
c.	*|A| + |A|	→	|A.A|

(5) Tier conflation



Fusion operations as in (4a–b) are asymmetrical in that one element acts as head and the other as operator. In KLV (1985) terms the operator only contributes its hot feature to the head's feature representation, resulting in one feature matrix. With the head represented on the right this implies that |A.I| and |I.A| are not equivalent thereby providing the possibility of defining more oppositions in a system based on three primitives. Apart from these three basic elements for vowels there were two further vowel components assumed. The so-called *cold vowel* plays the role of an identity element whenever it occurs in combination with other elements as operator because unlike full elements it does not have a hot feature. Its realization as head results in reduced vowels like schwa. The other vowel element was the |ATR| element which was argued as not derivable from the other elements and having the hot feature ATR.

Although elements were treated as feature matrices, the crucial point was that phonological processes have no access to features but only to the elements, with features manipulated only indirectly in element combinations. Elements within GP therefore directly address the overgeneration problem by positing an initially minimal set of primitives whose possible combination, even including head relations, fare much better on actual attested, in this case, vocalic systems.

A second issue that was deemed significant for sub-segmental units is their combinatorial properties and whether natural classes could be defined. In feature theory natural classes were later captured by assuming feature geometry (Clements 1985) by which particular nodes could be referenced as the target of particular phonological processes and by which segments can be modified. Combinatorial capabilities of elements in KLV (1985) were captured by charm, with elements regarded as either positively or negatively charmed. Elements of like charm repelled each other while elements of unlike charm attracted. This charm idea was adapted from particle physics where particles also operate on the same attraction principles. In this way elements could be grouped based on their charm and thereby capture natural classes by barring particular combinations while allowing others. In compound expressions the charm of the head determined the charm of the compound. In vowels relevant elements (A and ATR (also N)) were treated as positively charmed, while (I U v^o (the cold vowel)) were negatively charmed. Positively-charmed elements were cavity maximizing (oral |A|, pharyngeal |ATR| and nasal |N|), therefore allowing for high resonance and were as such regarded as unmarked so that markedness was defined both at the level of the phonological representation (read off the complexity of an expression) and at the elemental level. Thus to derive the unmarked minimal vowel system of {i u a} while adhering to elemental markedness, the ATR element was included as head to allow its positive charm to be propagated onto the phonological expression. An illustration is given in (6).

(6) Elemental expressions with charm
i (I−.ATR+)
u (U−.ATR+)
a (A+)

In this sense vocalic systems based on positively-charmed expressions are more unmarked than negatively-charmed ones with the obvious proviso that having only two negatively-charmed representations as in a system with {i u e ɛ ɔ o a} is less marked than having a full-blown negatively-charmed set of expressions as is attested in ATR systems. The incompatibility of (A+.ATR+) consisting of two positively-charmed elements accounts for the predominance of nine-vowel ATR systems rather than those with ten, and for the universal impossibility of having low +ATR vowels. The ten-vowel ATR systems required some further stipulations to be derived, further pointing to the markedness of an ATR /a/.

These representations suggested for vowels are extended to consonants in KLV (1990) with charm modified to having three values: positive, negative and neutral. KLV's (1990) focus was on defining government with charm and was also used to broadly refer to segments with assumed details of consonantal representations only discussed as far as they provided support for the theory of government proposed. The best articulation of these consonantal representations is given in Harris (1990) who further motivates the structure of consonantal expressions by a treatment of lenition processes as involving element decomposition in a bid to provide a non-arbitrary relation between phonological processes and the context in which they occur. The acoustic correlations of consonantal elements are developed in Harris & Lindsey (1995). The element set given in Harris (1990: 263) includes |I^{o}| ("^{o}" indicates neutral charm) as defining palatality; |U^{o}| as defining roundness in vowels and labiality in consonants; |v^{o}| denoting unmarked high and back attributes which contributes velarity when it is head in consonants; |N^{+}| as present in nasalized vowels and nasal consonants; |ʔ^{o}| as involving a decrease in overall amplitude achieved by a non-continuant gesture of the type that characterizes oral and nasal stops; |R^{o}| as correlated with a second formant transition that is characteristic of a coronal gesture; and |h^{o}| as a continuant characteristic treated as found in fricatives and approximants in KLV (1990) but which Harris (1990) treats as contributing a noise component in obstruents. The independent interpretability of elements is still assumed for consonantal elements: |ʔ^{o}| is independently interpreted as a glottal stop which contributes constriction in compound representations, and |R^{o}| is independently interpreted as a coronal tap. Thus fusion of |R^{o}| and |ʔ^{o}| produces a coronal non-continuant. The lack of any supralaryngeal gesture in |h^{o}| results in its independent interpretation as a glottal fricative. KLV (1990) propose |H−| and |L−| as source elements associated with stiff and slack vocal cords, respectively, defining non-spontaneous voicing in obstruents and tone on vowels. Neutral obstruents like in the Korean three-way obstruent system which have no active laryngeal gesture lack any source element. The full set of elements thus assumed is as given in (7) below.

(7) Set of elements (1990)

U^{o} - labial	R^{o} - coronal	N^{+} - nasal
I^{o} - palatal	ʔ^{o} - occluded	H^{-} - stiff vocal cords
v^{o} - none	h^{o} - noise	L^{-} - slack vocal cords
A^{+} - low		

In terms of distribution some elements are able to occur in both vocalic and consonantal positions while some are restricted to only consonantal positions, in particular |R^{o} ʔ^{o} h^{o}|. Charm theory also directly restricted the occurrence of elemental expressions (segments) whose charm value was determined by the head of the expression. Positively-charmed segments did not occur in consonantal positions while negatively-charmed segments did not occur in nuclear positions.

One motivation for these particular elements and their characterization is offered in the treatment of lenition given in Harris (1990). Harris treats lenition processes such as vocalization, spirantization and debuccalization as essentially involving the loss of complexity (or decomposition) in elemental compounds, where complexity is gauged as following from the number of elements that a particular phonological expression is composed of. Korean has a vocalization process affecting neutral stops that changes /p/ to /w/ and /t/ to /r/ which can in both cases be accounted for as loss of occlusion – element |ʔ^{o}| – leaving only the place-defining elements |U^{o}| and |R^{o}|, respectively. Such processes in Korean motivate the representation of neutral /p/ as |ʔ^{o}.U^{o}| and /t/ as |ʔ^{o}.R^{o}|. Treating lenition as depletion of elements makes the prediction that the susceptibility of a segment to lenition will be limited by the number of elements that the segment has. This therefore provides a means by which segmental composition can be determined. Thus in trying to account for the common diachronic lenition trajectory in (8a), the elemental representations in (8b) account for the process and thereby motivate the assumed representations.

(8) a. Common diachronic lenition trajectory:

plosive > fricative > h > ø

b.

```
x     x     x     x
|     |     |
Rº    Rº    |
|     |     |
hº    hº    hº
|
ʔº
[t]   [s]   [h]   ø
```

This approach to lenition, Harris argues, also specifically supports a monovalent approach to elements and their independent interpretability since it requires the pre-final position in a lenition trajectory to be interpretable. The approach also dispels any random substitution of features as may be permitted in a rule-based feature approach since the context of occurrence is non-arbitrary. The approach is also easily extendable to debuccalization as well as processes such as final obstruent devoicing and vowel raising which were previously considered as quite distinct from lenition. Harris further uses the notion of complexity as the basis on which governing relations can be determined, namely that governors are more complex than governees and it is probably fair to say that this led to the eventual demise of charm theory because it resulted in an incompatibility between markedness based on element complexity and inherent element markedness based on charm. This led to various different approaches to elements that aimed to address different issues specifically, but the central ideas that the approaches to be discussed below maintain is that elements are the central units of which segments are composed; they are monovalent and independently interpretable and employ some notion of headedness.

9.2.1.2 Phonetic interpretation and inventory size

The 1990s saw many changes in elemental representations in GP and extant element-based theories. With the loss of charm theory the idea of inherent natural classes was for the most part lost, and a number of approaches to elements aimed to replace this with another mechanism within a theory of segmental representations. One approach was to introduce tree dependencies akin to feature geometry, organizing elements in geometric representations (Harris & Lindsey 1995). This approach has remained in different forms in later element work (Kula 2002; Botma 2004; Nasukawa & Backley 2005). Another approach is to stipulate Licensing Constraints which monitor the combinatorial capabilities of elements (Charette & Göksel 1998).

The other issue of concern in the 1990s was the number of elements postulated and the inherent overgeneration of possible phonological units as opposed to actual attested segments. Related to the issue of overgeneration was the uncomfortable split between elements that only occur in nuclei and those that only occur in consonantal positions. Better parsimony would demand that all elements are able to occur in all positions, and one way of accounting for both overgeneration and distributional disparities was to dramatically reduce the number of primitives, leading to the so-called *Revised Element Theory* that uses between five and six elements (see e.g. Backley 2011; Charette & Göksel 1994/1996; Cobb 1997; Jensen 1994; Kaye 2001; Ritter 1997). A reduced set of elements is generally currently accepted and has led to a greater and different role of headedness in elemental representations (Backley 2011; Backley & Nasukawa 2009; Kula & Marten 1998; Nasukawa & Backley 2008) as well as the use of structural configurations to represent characteristics that may otherwise be subsumed by elements (Jensen 1994; Pöchtrager 2006). We will consider some of these issues in more detail below.

However, probably the most important change was a more precise characterization of elements as devoid of any direct association with articulatory properties. Starting in Kaye (1989) and further developed in Harris & Lindsey (1995), elements are defined as cognitive units fulfilling the grammatical role of encoding lexical contrasts, with the phonological component having a purely generative role of defining the grammaticality of phonological structures. The idea that elements and their corresponding phonological representations are characterized by full phonetic interpretation at all levels of derivation is the basis on which it is argued that there is no level of systematic phonetic representation. This follows from the fact that phonology does not involve the articulation and updating of abstract underlying representations which then have to be converted into physical phonetic objects since elements are always directly interpretable. Harris & Lindsey elaborate on how the phonetic exponence of elements are to be defined in acoustic terms, arguing that since the speech signal is the communicative experience shared by the speaker and the hearer, its primacy in phonetic interpretation cannot be in question. They propose acoustic signatures for elements as the patterns by reference to which listeners decode auditory input and speakers orchestrate and monitor their articulations (Harris & Lindsey 1995: 50). Backley (2011) provides further support and additional cues for these acoustic characteristics of elements. In Harris & Lindsey, where roughly the set {A I U (R) h ʔ ə (the replacement of v° above) (H L N)} are considered, the acoustic properties are partially as given in (7) above but with some refinements. The elements |H|, |L|, |N| were not tackled in Harris & Lindsey, and there is scepticism expressed on the validity of having an element (R) representing coronality, on which see discussion in section 9.2.1.2.2.

(9) Updated element acoustic signatures

A	–	central spectral energy mass (convergence of F1 and F2)
I	–	low F1 with high spectral peak (convergence of high F2 and F3)
U	–	low spectral peak (convergence of low F1 and F2)
h	–	noise manifested as aperiodic energy
ʔ	–	abrupt and sustained drop in overall amplitude
@	–	neutral spectral structure (non-coronal, non-palatal, non-labial, non-low)
(R	–	coronality)
H	–	aperiodicity
L	–	periodicity
N	–	nasality

An acoustic signature for |R| has proved elusive and, as will be discussed in section 9.2.1.2.2, this was one of the motivations that eventually led to the loss of |R| as an element in future developments. The elements when considered in consonantal representations can be divided into three types in line with segmental representations: resonance (or place), manner and source (laryngeal). As noted above, the resonance elements |A| |I| |U| define place for pharyngeals, palatals and labials, respectively. Pharyngeals are seen to result in lowering of vowels and so are associated with |A|, with the option that |A| may also be used to represent uvulars depending on the language inventory. See, for example, Bellem (2007) for a detailed analysis of emphatics employing resonance elements as central to characterizing different Arabic dialects. Palatalization occurring before front vowels and labialization before back vowels lend support to the resonance characterization of |I| and |U|, respectively. |ə|, without the inherent properties of the other resonance elements, can be used to represent velar resonance. The lenition processes, already discussed above, involving the stripping away of resonance properties support the characterization of |h| and |ʔ| as source elements. |N| together with |L| and |H| are best treated as laryngeal elements with properties that form the outer shell of a segment providing nasality or voicing. The status and combinatorial capabilities of these elements within segmental representations will depend on the contrasts expressed in particular languages since not all languages will exploit the full range of options offered by the system. This means that particular combinations of elements must be barred in particular languages in order to capture the natural classes the phonological processes form. Apart from tier conflation barring elements on fused tiers from co-occurring, other mechanisms are also used to achieve this effect.

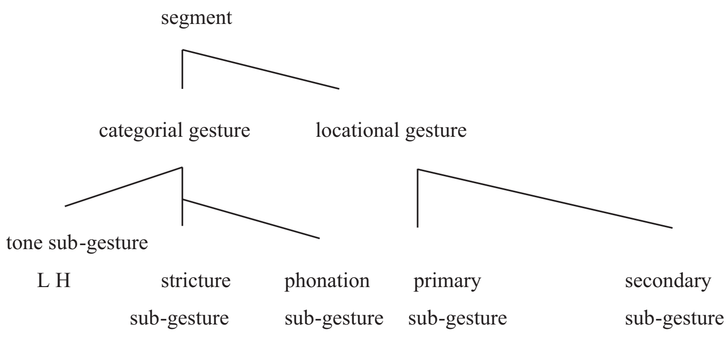
9.2.1.2.1 Natural classes: element geometries vs. Licensing Constraints

Elemental representations have also been concerned with the idea of natural classes and the ability of capturing the fact that particular phonological processes target specific sets of sounds to the exclusion of others. It remains a criticism of unary systems that it is not straightforward to, for example, refer to high vowels as a set since these do not share a characteristic element. An option is the less than ideal negative reference to those expressions that do not contain |A|.

In response to some of these challenges, including accounting for the fact that certain elements, e.g. place or source elements, do not routinely co-occur, feature geometry-like representations, akin to Clements (1985); Sagey (1986); and McCarthy (1988), among various others, have been attempted. Feature geometric representations allow reference to class nodes as a way of making reference to natural classes of sounds. There is a crucial difference, however, between feature and element geometries because, unlike features, elements do not require replacement of a deleted element in order to attain phonetic interpretability during the course of a phonological process because they enjoy independent interpretability. Thus while attaining phonetic interpretability is one of the intended outcomes of a feature geometry, this plays no role in an element geometry. An initial attempt at element geometries is given in Harris & Lindsey (1995) following the standard feature geometry groupings of laryngeal and resonance (place) nodes as class nodes containing the appropriate elements as terminal nodes. Manner elements attach directly to the root node. However, probably because an element geometry has more of an organizational function and the structure may be tacitly assumed, geometries have as such not played a very significant role in Element Theory. What it brings to the fore though is an explanation why certain combinations of elements are recurrent in representations of phonological expressions with a higher number of elements, and the connections that hold between different sets of elements. Thus rather than simply designating elements as being to do with manner, source or resonance – in the form of a statement – a geometry allows this to follow directly from a representation.

Geometries have also been used as a way of designating heads in representations in a principled way as motivated in the work of Smith (1988) and van der Hulst (1989), based on the premise that the same (combination of) sub-segmental units may acquire different interpretations depending on whether they are head or dependent. These ideas are further developed in Kula (2002) which aims to derive head relations by treating the characteristics which elements assume when they occur either as head or operator as following from their position in the geometry. Kula's (2002) geometry is based on Radical CV Phonology (van der Hulst 1994, 1995) which utilizes the notion of gestures as the organizing nodes of segmental structure. Under this view the segment is divided into two gestures: the categorial gesture and the locational gesture. The categorial gesture is further divided into three sub-gestures: the tone sub-gesture, the stricture sub-gesture and the phonation sub-gesture. On a par, the locational gesture is also divided into sub-gestures: primary location and secondary location sub-gestures. A particular quality of the assumed geometry is that the gestures and sub-gestures stand in a fixed head-dependent relation to each other: the categorial gesture is head of the locational gesture. Within the categorial gesture, stricture is head of the other two sub-gestures, and in the locational sub-gesture, primary location is head of secondary location. An illustration of a gesture-based head-dependent element geometry is given in (10) below where vertical lines identify heads.

(10) Head defining element geometry



A desired consequence of this element geometry with respect to the application of phonological processes in, for example, spreading processes, is that dependents will be able to spread independently, while heads must spread with their dependents. This, for example, explains the symmetry seen between the frequently spreading place features as opposed to the relatively stable stricture features.

In (10), the categorial gesture is chosen as the head of the whole segment because stricture distinctions generally determine the distribution of segments in syllabic organization. Within the categorial gesture, the representation of the tone sub-gesture as forming the outer shell of the categorial gesture characterizes its supra-segmental nature. The stricture sub-gesture contains elements that express different levels of stricture, such as absolute stricture |ʔ|, non-absolute stricture involving some interruption in unimpeded outflow of air |h| and unimpeded outflow of air |A| |I| |U|. The phonation sub-gesture expresses glottal stricture and voicing, viz. glottal stricture |ʔ|, glottal opening |H|, oral voice |L| and nasal voice |L|. The locational sub-gesture defines both consonantal and vocalic place articulations. The same place features are found in the primary and in the secondary sub-gestures with the difference that secondary place only occurs with some primary place specification.

Given this representation any elemental composition with a stricture element has that element as head and this defines the class of the segment as stop |ʔ| or fricative |h| or vowel |A| |I| |U|. If a stricture element is specified and the phonation element |L| is added, |L| is interpreted as voice. By contrast, if no stricture feature is specified and |L| is specified in phonation, it acts as head and has the interpretation of nasality. In this way we collapse |L| into concurrently interpreting voice and nasality, depending on whether it is head or dependent. This perspective thus interacts with the issue of element inventory size taken up below. This kind of geometry aims to provide a universal perspective on the combinatorial capabilities of elements, subsets of which are utilized in particular language inventories. Alternative representations that cover the entire consonantal spectrum are offered in Cyran (1997); Harris & Lindsey (1995); Rennison & Neubarth (2003); Scheer (1996, 1999, 2004); and van der Weijer (1996).

The other perspective adopted in GP to capture elemental co-occurrence restrictions within particular language systems is by so-called *Licensing Constraints* (LCs henceforth) which capture elemental combinatorial capabilities based on the phonological processes that occur in a language. See also other implementations of this idea in e.g. Dresher (2009) and Steriade (2007: 145f.). Their main purpose is to define the lexical set of elemental representations permitted in a language from a larger set of possible and well-formed expressions. Kaye (2001) provides some discussion of LCs. In a sense LCs are another way of expressing tier conflation although they differ in that they also include statements about headedness and licensing and thereby also constrain processes. There are usually different sets of LCs for nuclear expressions and for non-nuclear expressions although there is no requirement that these must be separate sets of constraints. Charette & Göksel (1998), for example, utilize LCs to define the vocalic system of Turkish and thereby also capture the vowel harmony process. Their set of LCs is given in (11).

(11) Turkish Licensing Constraints
  (i) Operators must be licensed
  (ii) |A| is not a licensor
  (iii) |U| must be head

LC (11i) requires all compound elemental representations to designate one of the elements as head where the head is deemed to license operators. Thus no expression should be headless. (11ii) means that A cannot be head in any compound expressions, while (11iii) requires |U| to be head in any expression in which it occurs. Given that there are three vocalic elements |A I U|, these constraints mean that the only other possible head in a phonological expression is |I| but that if |I| and |U| co-occur in the same expression then |U| will be head given LC (11iii). An identity element is assumed to be present in every expression with its empty content no longer represented by any symbol, having transitioned from |v| (the cold vowel) and |ə| (the neutral element). This results in the following licit Turkish nuclear expressions.

(12) Turkish nuclear expressions

a	|A|	e	|A.I|
i	|I|	o	|A.U|
u	|U|	ö	|A.I.U|
ü	|I.U|	ı	|_|

Heads in these expressions appear on the right and follow from the LCs in (11). The data in (13) illustrate two vowel harmony processes in Turkish that result in the alternation of the plural suffix *-lar* to *-ler* and the 2nd person possessive marker *-ın* to *-in/-un/-ın/-ün* in particular contexts.

(13) Turkish vowel harmony

stem		gloss	plural	2nd per possessive
a.	kil	'clay'	kil-ler	kil-in
b.	kül	'ash'	kül-ler	kül-ün
c.	kul	'subject'	kul-lar	kul-un
d.	kel	'bald patch'	kel-ler	kel-in
e.	köy	'village'	köy-ler	köy-ün
f.	kol	'arm'	kol-lar	kol-un
g.	kas	'muscle'	kas-lar	kas-ın
h.	kıl	'hair'	kıl-lar	kıl-ın

Charette & Göksel analyze vowel harmony in the plural as involving I-spreading from the stem nucleus to the suffix. All the plural forms that surface with *-lar* have no |I| in their stem vowel and so no harmony applies. In all other cases |I| spreads and is head in the resulting expression since according to LC(11ii) |A| cannot be a licensor. In the 2nd person possessive forms harmony involves |I| and |U| spreading which involves both elements spreading in forms where both are present in the stem vowel (13b,e). In resulting compound expressions |U| is head when it is present following LC(11iii). No spreading occurs when the stem does not contain a harmonizing element (13g,h). The point of significance is that the vowel harmony analysis proposed is not to be viewed as defined differently according to morphological processes but rather that the process globally involves both |I| and |U| spread. We do not see |U| spread in the plural because |U| only spreads into headless expressions, hence its spreading in the 2nd person possessive whose suffix is *-ın* with no element in the suffix vowel. The option to switch heads after |U| spreading is one not exploited by Turkish but which applies in other Turkic languages (Charette & Göksel 1994).

LCs for consonants are discussed and illustrated in Kula & Marten (1998, 2000) and Kula (2005) for some Bantu languages (Bemba and Herero) showing how the LCs proposed also feed into the analysis of assimilation and strengthening in NCs and nasal consonant harmony. LCs therefore define language particular systems and provide a means by which phonological processes are derivable.

9.2.1.2.2 Inventory size: expansion vs. reduction

One of the all-time consuming issues in GP and Element Theory is the number of basic primitives that should be assumed. There has been a concern to control the expressive power of the theory to more approximately match the 100 or so speech sounds attested in the world's languages. There have been proposals to reduce the number of elements from around twelve primitive elements {A I U R h v/ə ATR ʔ N L H} to five or six {A I U ʔ H L}, with the choice of six seemingly gaining ground. While the avoidance of overgeneration is undoubtedly to be admired, this comes with a number of consequences to ensure that the limited set can still generate the requisite number of contrasts and account for the attested phonological processes.

At least all proposals have lost the neutral elements (ə, v) which were in any case always considered to be the representation of emptiness. The idea that there is an identity element or empty element remains but gets no symbol in order to more appropriately highlight that it is empty. Thus empty nuclear positions if they are phonetically interpreted get realized by a sound (schwa, ɨ, ı, etc.) that has no elemental content. The ATR element has also been lost mainly owing to the fact that its use was restricted to vocalic positions and its validity could not quite be extended to vowel systems that do not have an ATR contrast. In addition, as Harris (1990) argues, an ATR element results in an asymmetrical representation of the sets of ATR vowels vs. their non-ATR counterparts with the former always being more complex. Harris chooses to replace the ATR element with the neutral element |ə| understood as the acoustic baseline or canvas on which elemental patterns are superimposed, independently interpreted as schwa. In representations where |ə| is operator it contributes nothing as a neutral element and when it is head in vowels the elemental characteristics of |A| |I| |U| are somewhat backgrounded in operator role giving the phonetic output of non-ATR. This way the representation of ATR and non-ATR vowels are identical in terms of elemental composition/complexity and differ only in which element is head. In this sense ATR harmony does not involve the spreading of features but agreement in head, i.e. non-ATR vowels must all share |ə| as head so that all vowels in a harmonic span are head aligned. The same holds for the ATR counterparts which also share the element that is head in their harmonic span assuming that headship is assigned on the autosegmental tier of the element in question.

This idea of ATR harmony being captured by head alignment is developed further as the sole means of expressing ATR harmony after |ə| is discounted as an element (see e.g. Walker 1995). Under this revised approach where |ə| is treated as no longer part of the element set, ATR systems are represented in terms of headedness with the ATR set being all headless while the non-ATR set is all headed thereby still retaining a balance in the complexity of the two sets. Harmony is achieved as head-licensing where the harmonic span of a word is defined by the propagation of headedness to all nuclei in the domain emanating from the domain head. In this way morphological alternations triggering agreement in ATR are treated as agreement in head status achieved via head-licensing. Examples of ATR harmony in many West African languages that can be characterized in this way abound in the literature: Vata (Kaye 1982), Akan (Clements 1981), Yoruba (Archangeli & Pulleyblank 1989), among others.

The other element that has been strongly called into question is |R| representing coronality. Apart from the fact that a clear acoustic signature for |R| has remained elusive, Harris &

Lindsey (1995) argue that treating coronality as represented by |R| fails to reflect the special distributional attributes, behavioural characteristics and uniqueness of coronals as has been shown in Paradis & Prunet (1991) and much subsequent work. Backley (1993) also strongly endorses the loss of |R|, suggesting that perhaps coronals are placeless owing to their susceptibility to assimilation in contrast to other places of assimilation, including the fact that they behave transparently with respect to a number of processes. These concerns have resulted in the diminished use of |R| with coronality derived in different ways according to its distributional properties in different language systems.

Thus with the loss of {ə v ATR R} the elements {A I U h ʔ L N H} remain. Within this set |h| and |H|, on the one hand, and |L| and |N|, on the other, are merged so that currently the most dominant set of elements employed is one that assumes six elements viz. {A I U ʔ L H} in what is sometimes referred to as the revised set of elements (see e.g. Backley 2011; Charette & Göksel 1994/1996; Cobb 1997; Jensen 1994; Kaye 2001; Ritter 1997). Recall that |h| represented frication and |H| voicelessness while |L| represented voicing and |N| nasality. The merger of these pairs of elements is not to assume that the two properties concerned are indistinct but rather that there is a strong link between them which is considered as best captured as a difference in headedness. Thus |H| as head in a phonological expression represents frication but voicelessness as operator, and in the same vein |L| as head represents nasality but voicing as operator. A number of phonological processes that show a relation between each of these pairs of properties gain significant insight by this assumption, indeed motivate this representation. A number of studies have argued for the unification of voicing and nasality in the literature (e.g. Botma 2004, 2009; Botma & Smith 2006, 2007; Botma et al. 2013; Kula & Marten 1998; Kula 1999, 2002; Nasukawa 1995, 1998, 2005a; Ploch 1999).

Nasukawa (1997, 2005a), for example, considers the well-known facts of voicing in Yamato Japanese NC clusters and Rendaku (Itô & Mester 1986). There are two relevant sets of data: The first concerns post-nasal voicing and the second cases of (Rendaku interacting with) Lyman's law, which prohibits two voiced consonants in a domain. Consider the examples below.

(14) Japanese post-nasal voicing and rendaku

a.	shombori		'discouraged'	*shompori
b.	shindoi		'tired'	*shintoi
c.	kaŋgae		'thought'	*kaŋkae
d.	shin+te	γ shinde	'die' (gerundive)	
e.	kam+te	γ kande	'chew' (gerundive)	
f.	onna+kokoro	γ onnagokoro	'woman's heart'	
g.	kami+kaze	γ kamikaze	'divine wind'	
h.	ori+kami	γ origami	'paper folding'	

(14) shows that voiceless obstruents get voiced when preceded by a nasal (14a–e), thereby apparently indicating that the nasal has a "voice" element which spreads. On the other hand, the Rendaku data in (14f–h) contradict the idea of a "voice" element present in nasals since Rendaku, which requires the initial consonant of the second member of a compound to be voiced, still applies in apparent violation of Lyman's law if nasals are treated as bearing a "voice" element (14h). This apparent contradiction can be accounted for, Nasukawa argues, by treating "nasality" and "voicing" as reflexes of the same underlying element |L|, depending on whether it is head (voice) or operator (nasality).[4] From this perspective, the Rendaku facts can be explained as reflecting the fact that two |L|-headed expressions are disallowed

in the second member of a compound. Since the |L| element in nasals is operator it does not violate Lyman's law. Post-nasal voicing, on the other hand, is accounted for as rightward |L| spreading into adjacent onsets in a particular configuration, here onset-to-onset government. The spreading requires a change in headedness to achieve the voicing effect. The derivation of *shinde* "die (gerundive)" from the base *shin* plus the genitive suffix *-te* is as shown below (Nasukawa 1997: 418). Elements that are head are underscored.

(15) Japanese post-nasal voicing: /shin-te/ γ [shinde]

```
O N O N O N    →    O N O N O N
| | | | | |         | | | | | |
x x x x x x         x x x x x x
| | |   | |         | | |   | |
ʃ i n   t e         ʃ i n   d e
    [ʔ] [ʔ]             [ʔ] [ʔ]
        [h]                 [h]
    [L]                 [L] → [L]
```

(15) shows the rightward spread of |L|. |L| cannot remain as operator in the recessive Onset because in Nasukawa's analysis there is a constraint against the co-occurrence of element |h| and an operator element |L|, formalized in terms of tier fusion, namely, the two elements reside on the same tier. From this configuration we expect NC clusters to behave like voiced consonants in Rendaku since they bear an |L| head. Note that the change in headship could be accounted for along similar lines as Lyman's law, i.e. as an OCP-motivated constraint against two adjacent |L| operators. Needless to say the dual role associated with one element – here |L| – depending on whether it is head or not elegantly captures the interaction between nasality and voicing without the need of employing a voice specification in spontaneously voiced nasals.

Further merger in a way is expressed in the representation of tone. The laryngeal-source elements |L| and |H| are also used to represent tone based on the empirical evidence showing interaction between these elements and tone. In contrast to voicing, which is regarded as only occurring in consonantal positions, tone occurs in vocalic positions, in which case syllabic nasals that may bear tone are represented as partially occurring in nuclear positions. Used as tone elements the autosegmental nature of tone is captured by representing such elements on tiers as discussed above (see Kula 2012 for some discussion on tonal representations). There are a number of processes that show an interaction between tone and voicing that support the assumption that the two are represented by the same element. The classic example is depressor consonants involving the lowering of a high tone by a voiced consonant. Discussion can be found in Trail et al. (1987) for Zulu, Bradshaw (2003) for SiSwati, and Pearce (2009) for Kera, for example. The examples from Zulu in (16) below show a two-way contrast of high and low tone in its tone system. Voiced consonants are seen to depress (i.e. lower) the tone of the syllable in which they occur. High tone is indicated by an acute accent and low tone by a grave accent. In each of the paired examples showing a singular and a plural the nominals have an initial VCV noun class marker which has two high tones (HH) in the singular but a HL structure in the plural which contains a voiced consonant /z/.

(16) Zulu depressor consonants

a.	ísí-khwámà	'bag'
	ízì-khwámà	'bags'
b.	ísí-hlálò	'seat'
	ízì-hlálò	'seats'
c.	ísí-fúndò	'lesson'
	ízì-fúndò	'lessons'
d.	ísí-kòlè	'school'
	ízì-kòlè	'schools'

These facts are explained in Element Theory under the view that depressor consonants contain |L| and that this laryngeal specification must hold over the minimal licensing domain of onset–nucleus (CV pair). This implies that within the voicing specification of Zulu, voiceless sounds (as well as sonorants) are unspecified for a laryngeal feature while depressor consonants are specified with |L|. Since this specification causes a clash with a high tone, changing it to low, we must conclude that voicing and low tone are connected and must therefore be represented with the same primitives. In this case the voice specification of the depressor |L| is interpreted as low tone on the adjacent vowel. This thus demonstrates the full extent of oppositions for |L| as either voice or nasality in consonants and as either tone or nasality in vowels, with headedness used to create a contrast in each position. We return to the matter of headedness presently.

By comparison there has been much less discussion of the fusion of |H| and |h| although this merger has been assumed in element models adopting a maximum of six elements (but see the later discussion of Backley 2011 for some motivation). An alternative view on the status of |h| is offered in Cyran (1997, 2010) who argues for a parametric perspective on the presence or absence of |h| in a particular language system. This assumption allows for a more nuanced representation of voiced fricatives as following from the contrasts expressed in a particular language system in this sense contrasting Irish and Polish, for example. Thus while |h| may be considered to be parametrically present in Polish, which has both voiced fricatives and affricates, |h| is considered as not present at all in Irish, which has no voiced fricatives or affricates. This highlights an important issue in elemental representations, namely that different compositions of elements can result in the same phonetic object with the crucial factor being that representations are motivated by phonological patterning in particular languages.

Thus there remains two major sets of elements assumed in GP. One that is considered more traditional because it retains the use of |h| and possibly |R| in the set {A I U ʔ h (R) H L} and the reduced version that uses only six primes: {A I U ʔ H L}.

9.2.2 Issues and directions

Within sub-segmental representation a number of issues form part of current discussion in GP, extending the theory into new territory as well as reconsidering and refining issues that are central to the theory. We will consider three issues here: headedness, complexity and the nature of primes.

9.2.2.1 Headedness

The representation of segmental contrasts faces a number of challenges under the Revised Element Theory which currently assumes maximally six elements |A I U ʔ H L|. This has led to a revision of some previously held assumptions. The greatest challenge is being able to express a sufficient number of contrasts while also maintaining the assumptions on independent interpretability of elements and treating each element as having its own identifiable signature as assumed in Harris & Lindsey (1995). In this sense headedness has always played an important role in the expressive power of Element Theory (KLV 1985). In complex expressions it signals that the characteristic of the element which is head is more enhanced in an expression. This, for example, captures the contrast between tense and lax mid vowels which are otherwise considered as containing the same elements. Under this perspective a complex expression only has one head, but see section 9.2.2.2 for possible alternative views.

We see headedness already beginning to take on a different role in the late 1990s in standard GP where Ritter (1997), for example, argues to replace stricture, in particular stopness expressed by |ʔ|, with headedness. In this case a stop like /p/, which would be represented as |ʔ.h.U|, gets represented as |h.$\underline{U}$| with headedness used to contrast it from the fricative /f/, which has identical elements |h.U|. This move not only redefines headedness but also further reduces the element set by dispensing with |ʔ|. In more recent work we see further redefinition of headedness as, for example, presented in Backley (2011) with precursors in Nasukawa & Backley (2005) and Backley & Nasukawa (2009). In this approach headedness is no longer viewed as the enhancement of a particular elemental characteristic within a complex expression but is rather equated to the identification of an independent acoustic signature identifying a distinct characteristic from its non-head counterpart. Velars, which often receive varying representations in different elemental analyses, are treated as represented, like labials, by |U| with only a contrast in headedness between the two places of articulation.[5] The symmetrical treatment of velars and labials echoes Jakobson et al.'s (1952) use of the acoustic feature [grave]. This representation is supported by the unity seen between velars and labials in different phonological processes attested in a number of languages. This observation of the affinity seen between [u, w] and both labials and velars has also been noted in Scheer (1999) who opts to represent labial and velar place with different elements owing to the fact that roundness (labiality) does not always go hand in hand with velarity. The intuition of this analysis is captured in Backley's (2011) representation of velars with |U| and labials with |$\underline{U}$| with the consequence that headedness is no longer viewed as the greater contribution of a consistent phonetic characteristic of an element to its expression. While the interaction of [u, w] with labial consonants is well attested, there are similarly a number of examples demonstrating interaction with velars. Consider the following two examples from Moroccan Arabic and Czech discussed in Scheer (1999: 209). (17) shows broken plural formation in Moroccan Arabic where only velar and uvular consonants allow labial secondary articulation, whereas other places of articulation do not, so that *[s^w, d^w], for example, are unacceptable. In broken plural formation a labial [w] targets the initial consonant of the root.

(17) a. Labial secondary articulation possible

singular	broken plural	
kbir	k^wbar	'tall'
χurza	$χ^w$razi	'node'
quamiʒa	q^wamɨʒ	'shirt'

b. Labial secondary articulation impossible

singular	broken plural		
amin	sman	*s^wman	'fat'
silla	slali	*s^wlali	'basket'
drif	draf	*d^wraf	'nice'

In Czech in (18) below we see a similar compatibility between [u] and velars in the distribution of three vocative allomorphs in consonant-final masculine nouns. *-i* occurs with palatals, *-u* with velars and *-e* elsewhere.

(18)		nominative		vocative
	a.	kuuɲ	kɔɲ-i	'horse'
		tomaaʃ	tomaaʃ-i	'Thomas'
		złɔɟɛj	złɔɟɛj-i	'liar'
	b.	hox	hox-u	'boy'
		zdɛɲɛk	zdɛɲk-u	*a given name*
		ptaak	ptaak-u	'bird'
	c.	pɛs	ps-ɛ	'dog'
		dɔktɔr	dɔktɔr-ɛ	'doctor'
		ɦɔlup	ɦɔlub-ɛ	'pigeon'

Both examples show a compatibility between velars and [u,w] that supports a segmental representation that captures these distributional facts. Backley's (2011) representation of these two contrasting places of articulation as a difference in headedness suggests a reinterpretation of headedness. However, he continues to assume an asymmetry between headed and non-headed expressions, arguing that non-heads are more likely to be the target of assimilation processes and to occur in weak positions. He then treats this as the basis for deciding which of the two segment types (velars or labials) will be treated as having headed |U̲|. Examples that support the representation of velars as non-head are Selayarese (Malayo-Polynesian) where velar resonance may be over-ridden by another resonance property in reduplication processes where the place of a final velar nasal is changed to that of the following consonant of the reduplicant. Another example is given by Skikun, an Atayalic dialect of Formosan (Northern Taiwan) where Backley presents data showing an ongoing change in young speakers where labials are changing to velars only in coda position. Final /p, m/ are changed to their velar counterparts, reflecting that the change from labial to velar consists of a weakening process. In this sense headedness can be used to represent phonological weakness and strength, with non-headed expressions more likely to occur in weak positions such as the coda. A similar argument is presented for coronals and palatals which are both represented with |I| as their resonance element, the former being non-head and the latter being headed. In this case as well, the choice of coronals as being non-head follows from their susceptibility to phonological processes as is cross-linguistically attested and noted in earlier discussion. Thus the choice of head is not random but connected to phonological patterning (see Backley 2011: 72f. for some examples). Since the headed and non-headed counterparts of an element are argued to have their own independent spectral patterns, though they share a central characteristic, they are able to co-occur in, for example, doubly articulated segments such as the labio-velar approximant /w/. Since all elements are able to occur in both nuclear and Onset positions but take on slightly different characteristics when they are headed or not, every element has four different possible interpretations, two each in Onset and nuclear position. In this way headedness plays an important role in expanding possible contrasts while maintaining a relatively small set of primes. Note though that importantly headedness is restricted to resonance elements (as also argued in Scheer 1999). Backley's line of argumentation thus allows for the possibility that the "same" element may recur in a phonological expression and potentially opens the door to doubly headed expressions, otherwise barred in GP representations (though see Rennison & Neubarth 2003).

In a slightly different way source elements can also be seen to contribute different characteristics depending on the segment type in which they occur and importantly in the kind of language system in which they occur. Voicing contrasts are one classic example that demonstrates this variable interpretability of source elements. This is depicted for |H, L| in (19) below from Backley (2002: 8), providing a summary of the possible interpretations of |H| and |L| in generating segmental contrasts. Similar representations are argued for in Nasukawa (2005b); Cyran (2010, 2014); and Botma et al. (2013), among others.

(19) a. |H| : high tone on vowels (e.g. languages with lexical tone)
|H| : aspiration in plosives (e.g. English, Korean)
|H| : audible friction in fricatives (e.g. most languages)
b. |L| : low tone on vowels (e.g. languages with lexical tone)
|L| : nasality in vowels and sonorants (e.g. most languages)
|L| : full voicing in obstruents (e.g. French, Japanese)
c. |H,L|: contour tones on vowels (e.g. languages with lexical tone)
|H,L|: breathy voicing in plosives (e.g. Gujarati)

This representation of voicing aptly captures the distribution of VOT cross-linguistically. As presented in Cyran (2010, 2014), languages without a voicing contrast have no voicing specification, those with lead VOT have only |L|, while those with VOT lag have only |H|. Systems with both VOT types imply a three-way contrast and languages like Hindi and Gujarati with a four-way contrast imply a combination of the two elements as (20) below shows with some example languages drawn from Cyran (2010: 16).

(20) Voicing contrasts captured with |H| and |L|

<table>
<tr><th>language</th><th colspan="3">VOT opposition</th><th>representation</th><th>examples</th></tr>
<tr><td>Malakmalak</td><td></td><td>–</td><td></td><td>|_|</td><td>p</td></tr>
<tr><td>Spanish, Polish</td><td>lead</td><td>–</td><td></td><td>|L| |_|</td><td>b, p</td></tr>
<tr><td>English, Irish</td><td></td><td>–</td><td>lag</td><td>|_| |H|</td><td>b, p^h</td></tr>
<tr><td>Thai</td><td>lead</td><td>–</td><td>lag</td><td>|L| |_| |H|</td><td>b, p, p^h</td></tr>
<tr><td>Hindi</td><td>lead</td><td>–</td><td>lag, lead/lag</td><td>|L| |_| |H| |LH|</td><td>b, p, p^h, $b^ɦ$</td></tr>
</table>

An underscore is here used to represent absence of an element. Notice that with the use of underspecification, headedness is not invoked at all in generating voicing contrasts.

9.2.2.2 Complexity

The discussion of non-headed expressions occurring in weak positions and their counterpart headed expressions in strong positions raises the important issue of complexity and the representation of sonority effects in Element Theory which continues to be an important issue in current work. Sonority is generally argued to be responsible for the structure of the syllable: Onset clusters with rising sonority are considered to be unmarked because they maintain the desirable rising sonority profile for a syllable. From the GP perspective this adoption of an observation as explanatory is unsatisfactory because it does not follow from any general principles.

Recent GP work offers at least two possible ways of capturing the robust cross-linguistic patterns that are attributed to sonority. In both approaches it is essential that sonority effects follow from not only the internal structure of segments but also from the accompanying government and licensing relations that the segments are engaged in. Both sub-segmental and segmental relations derive complexity captured as substantive vs. formal complexity, respectively, in the approach adopted by Cyran (2010). In this perspective substantive complexity is the inverse of sonority and strength, following the basic assumption in Harris (1990) which captures complexity as transparently reflected in the number of elements of which a segment is composed. Obstruents with a higher number of elements are the most complex and therefore, from a standard GP perspective, act as governors of less complex liquids both within branching onsets (left-headed) and in coda–onset clusters (right-headed) (see Chapter 10, section 2).

Note that the notion of complexity is only available in an environment where primes are privative: when features are assumed, all segmental expressions are (or, in underspecification approaches, end up being) made of the same number of primes. It is only when primes may be either present or absent that a different number of primes can be responsible for their makeup.

Harris' work as discussed earlier infers the higher complexity of obstruents as following from lenition processes which are accounted for as the loss of segmental complexity (see the decreasing complexity of the lenition trajectory t → s → h under (8)). In this sense the strength attributed to stops is transparently captured by their greater sub-segmental complexity. Cyran (2010) argues that this complexity pattern (in conjunction with formal complexity concerning syllabic patterns) is the basis on which language inventories, phonotactics, typology, markedness and phonological processing are organized and can be explained. An important addition is that substantive complexity is treated as scalar, providing a non-arbitrary scale with cut-off points, although exactly where the grammar of a particular language chooses to place the divisions remains an arbitrary property of each grammatical system. In this way Cyran provides an analysis employing different complexity scales to define the segmental inventories of English, Polish and Irish showing how these different languages manipulate complexity in sub-segmental structures in a fashion that is most economical for each system and depending on the contrasts expressed in each system. What is crucial is that sonority effects such as syllable contact and phonotactics are still derivable from complexity in each system.

An alternative view is offered in Scheer (1999, 2004) under the consideration of the distributional properties of word-initial clusters. The approach adopted in this case also similarly argues that sonority effects follow from a number of general principles including Government Licensing (Charette 1990), the ECP (KLV 1990) and the initial empty CV (Lowenstamm 1999; Scheer 2004), rather than just the complexity of segments. On segmental complexity and adopting a strict CV approach, Scheer argues that for the purpose of the grouping of segments into Onset clusters only resonance elements count (A, I, U). Evidence from regular segmental alternations then shows that liquids (and sonorants more generally) possess more resonance elements than obstruents and are therefore governors. In Strict CV, Onset clusters (TøR) as much as coda–onset clusters (RøT) are separated by an empty nucleus – they are distinct because the empty nucleus enclosed in coda–onset clusters needs to be governed, while the one separating Onset clusters is circumscribed by a sonority-based relationship whereby the sonorant governs the obstruent. This relationship is called infrasegmental government (see Chapter 10, section 2.3, and Chapter 11, section 1.3). Hence the consonants of a TøR cluster do interact, while those of an RøT cluster do not. The fact that R governs T in the former (rather than the reverse) follows from Government Licensing (governors need to be licensed by the following expressed nucleus: in TøRV, the empty nucleus cannot government-license the T, but the V can license the R) and the basic principle of standard GP that relations among constituents are head-final (in Strict CV constituents do not branch, hence there are no intra-constituent relations). This setup (as well as the initial empty CV site) captures the cross-linguistic distribution of word-initial consonant clusters as discussed in more detail in Chapter 10, section 1.7.

Finally, note that both approaches to complexity were worked out in the regular ten-element system of standard GP, i.e. before the reduction of this set as discussed above was undertaken. Reduction concerns the elimination of some elements (R, ATR, ə, h, N, eventually ʔ) and multiple function of others according to their syllabic affiliation and headedness (L, H). Since this concerns only non-resonance elements (with the exception of R), that is precisely those which make obstruents more complex than sonorants in Harris' approach, the complexity debate may need to be reassessed. Complexity being the pivot articulating syllable structure and melody, though, it is theory-specific and hence dialectically interleaved with specific assumptions on syllable structure.

It was mentioned that complexity is a fundamentally different approach to sonority than what is found in (all) other theories: only privative primes produce segmental expressions that are made up of a contrasting number of primes. As a consequence, there are no specific primes encoding sonority: items such as [±son], [±cons], [±voc] etc. do not exist in GP – their function is taken over by complexity. This means that sonority is not melody but a function computed upon melodic makeups (counting the number of primes). Therefore GP predicts that whenever melodic primes behave like a class and are opposed to non-melodic phonological properties, sonority will not behave like a piece of melody. This is reflected by the fact that sonority (and only sonority) may be read off regular autosegmental representations: sonority is the only "melodic" property that is projected at the syllabic level. Indeed, encountering a branching Onset reveals the relative sonority of the segments involved, and the same is true for coda–onset clusters. By contrast, syllable structure alone does not provide any hint as to whether the segments at hand are labial, dental or velar, voiced or voiceless etc. Hence sonority is visible from above the skeleton, while place and laryngeal properties are not.

That the non-melodic nature of sonority may be on the right track is shown by the fact that there are a number of phenomena that cannot take any melodic properties into account, but are sensitive to sonority. Phonologically conditioned allomorphy is a case in point (Scheer 2016), and so appear to be crazy rules (which are only ever segmentally crazy; Scheer 2015), category-sensitive phonology (i.e. specific phonologies applying to nouns and verbs; Smith 2011) and absolute agrammaticality (which appears to be only due to non-melodic properties).

That sonority is ontologically different from melody is also suggested by vocalic sonority. Summarizing his typological work on vocalic properties that influence stress placement, de Lacy (2002: 93) formulates the following amazement.

> One issue this typology raises is not why stress is sensitive to sonority, but rather why it is not sensitive to so many other properties. There are no stress systems in which subsegmental features such as Place of Articulation or backness in vowels plays a role in assigning stress. The same goes for features such as [round], [nasal], and secondary articulation.

This touches upon the issue of vocalic sonority: as it stands, complexity makes the correct predictions regarding the difference between high and mid vowels. The former are made of one prime, while two primes contribute to the latter (in the case of front rounded vowels the contrast is between two and three primes). That is, more sonorous vowels are more complex (which parallels Scheer's approach to consonantal sonority where also the more sonorous items are more complex). The typical low vowel /a/, however, is made of just one element and therefore fails to express this bit of vocalic sonority in terms of complexity.

9.2.2.3 Structural representation of primes

The number of elements used in GP and Element Theory continues to be an issue of debate even though the use of six primes has steadily gained ground. The reduction to six primes predicts 256 possible segments if all permutations are allowed (Jensen 1994). This has generated debate that an even smaller number of primes may be preferable, with five elements generating the more realistic 112 possible segments (assuming only one element can be head in any expression). Work in the early 1990s (see Jensen 1994) was particularly concerned with this notion of reduction of primes. Jensen (1994) thus argues that phonological differences which have been attributed to the presence of the glottal element |ʔ| can in fact be expressed by differences in constituent structure. It is particularly argued that the "stop" impression is the reflex of a branching rhyme structure preceding the relevant assumed |ʔ| containing phonological expression. The absence of such an environment results in the interpretation of the same phonological expression with a "fricative" impression. In this way |ʔ| would not need to be expressed as an independent element. This idea of replacing an element with structural configurations has not been much pursued in the GP/Element Theory literature, but recent work of Pöchtrager (2006); Pöchtrager & Kaye (2013); and Živanović & Pöchtrager (2010) in a version of GP they loosely term GP 2.0 take up this line of argumentation. The idea is to utilize hierarchical structural configurations (akin to minimalist syntax) to capture the properties of the elements |ʔ H A| which are argued to either show no generalizability to consonantal and vocalic positions or to display special characteristics that justify an enriched structural configuration. This implies that the element set is limited to the three remaining elements: |I U L|. This approach also abandons an autonomous melodic tier so that elements are now annotations on terminal nodes and are thus embedded in phonological structure. The role of melody is thus depleted with phonological phenomena previously treated as melodic now considered as structural. A discussion of the inventory and a number of phonological properties of Putonghua is offered in Živanović & Pöchtrager (2010), providing an illustration of the main tenets of this approach still in its infancy.

9.3 Conclusion

Element Theory within GP remains a competitive theory of sub-segmental structure. Like any theory it has undergone a number of changes, resulting in slightly different versions in current operation. The properties of privativity, unarity, cognitive import and independent interpretation remain representative of this approach. What makes the approach distinctive from other theories of melodic structure is the unary character of primes, i.e. their size. Indeed, |A I U| etc. are bigger than a single feature since they describe articulations that are made of several features (I for example is high, front and unrounded). Their size is also the reason why elements may be independently pronounced. Finally, elements are not based on articulation and involve a minimal number of distinctions. A reduced set of primes on grounds of avoidance of overgeneration is preferred and deemed theoretically more desirable. Overall much more work has been conducted on the representation of vowels, and the use of |A I U| as resonance elements has received much cross-linguistic investigation and exemplification, although the representation of resonance in coronals remains contentious. Work on consonantal representations, though present, would benefit from more systematic investigation across a larger set of languages. Although universal phonetic cues are associated to each independent element, the phonological composition of a segment within a language remains motivated by the phonological processes therein, with the phonetic interpretation only becoming fully meaningful when viewed as part of a sound system.

9.4 Further reading

Kaye, Jonathan 1989. *Phonology: A Cognitive View*. Hillsdale: Erlbaum.
Conceptual underpinnings of Government Phonology: why there is phonology at all, the relationship with phonetics, the organization of the lexicon, parsing cues in perception.

Harris, John & Geoff Lindsey 1995. The elements of phonological representation. *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 34–79. Harlow, Essex: Longman.
Interface with phonetics: interpretational autonomy of phonology, no underspecification, no level of systematic phonetic representation. The output of phonological computation is directly converted into phonetic values through a dictionary-type conversion (phonetic interpretation).

Scheer, Tobias 2015. How diachronic is synchronic grammar? Crazy rules, regularity and naturalness. *The Handbook of Historical Phonology*, edited by Patrick Honeybone & Joseph C. Salmons, 313–336. Oxford: Oxford University Press.
Small is beautiful vs. big is beautiful: how much of the pool of surface alternations are phonological in kind?

Gussmann, Edmund 2007. *The Phonology of Polish*. Oxford: Oxford University Press.
How computation works in standard GP: no extrinsic rule ordering, but rather unviolable, unweighted constraints. Detailed application to palatalization in Polish (chapter 3).

Backley, Phillip 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.
The book offers a recent take on elements and important developments since the early work. Specific proposals on what role headedness can play in extending element sets is offered. A wide range of languages are investigated.

Cyran, Eugeniusz. 1997. *Resonance Elements in Phonology: A Study in Munster Irish*. Lublin: Folium.
A study that provides a good grasp of the central elements |A| |I| |U| not only as they are used in vowels but also in consonants drawing on the parallelism between the two segments types. Clear specific analysis and how they may vary are offered.

Nasukawa, Kuniya. 2005. *A Unified Approach to Nasality and Voicing*. Berlin and New York: Mouton de Gruyter.
The book discusses in the detail the option of merging some elements to have a dual identity, here with respect to nasality and voicing. A number of phenomena that support such a view are discussed.

Backley, Phillip & Kuniya Nasukawa 2009. Headship as melodic strength. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 47–77. Berlin and New York: Mouton de Gruyter.
This work takes up the issue of headship to show how it can be exploited to lead to specific interpretations of elements, with those elements that are head argued to acoustically have a stronger signature and by virtue of which they must make specific contributions.

Kula, N. C. 2012. On the representation of tone in Element Theory. *Sound Structure and Sense*, edited by Cyran et al., 353–367. Lublin: Wydawnictwo KUL.
An account of how the elements |L| |H|, which are also associated with tone, play a role in this autosegmental property of segments and how this interacts with voicing in consonants to result in depressor effects in a number of Bantu languages.

Botma, Bert, Nancy C. Kula & Kuniya Nasukawa 2013. Features. *Bloomsbury Companion to Phonology*, edited by Nancy C. Kula, Bert Botma & Kuniya Nasukawa, 33–63. London: Bloomsbury.
Gives more background on the difference between feature systems and element-based approaches and how these compare. A case study on the interaction of voicing and nasality is offered on Zoque.

Scheer, Tobias 1999. A theory of consonantal interaction. *Folia Linguistica* 32: 201–237.
This paper is gives further exposition on how relations above the segment may be affected by the elemental make up of segments, defining infrasegmental government. Some detailed discussion of elemental representations is offered.

Notes

1 Cases of this kind of phonology–phonetics mismatch are quite frequent. For a parallel case from Japanese where [ɯ] represents both |U| and an empty nucleus, see Nasukawa (2010).
2 See Bermúdez-Otero (2007) and Scheer (2015) for a comparison with other theories and the relationship with diachronic development.
3 With one exception applying to so-called analytic morphology, which is subject to a no-look-back device that in current syntax is called the Phase Impenetrability Condition (on which more in Chapter 11.1.4).
4 The choice between whether |L|-head instantiates voicing or nasality differs between different researchers reflecting that what matters is the opposition. Indeed, the use of |N| or |L| as the symbol for the merged element representing voice/nasality also varies. These differences must be treated as entirely superficial.
5 A number of analyses treat velar as empty as in e.g. Huber (2003), among others. This analysis has precursors in early Element Theory where velar was represented by the cold vowel |v| as in e.g. Harris (1990). See also Szigetvári (1994) on both the representation of velars and coronals.

References

Anderson, John M. 1985. Structural analogy and dependency phonology. *Acta Linguistica Hafniensia* 19: 5–44.

Anderson, John M. 1992. *Linguistic Representation: Structural Analogy and Stratification*. Berlin & New York: Mouton de Gruyter.

Anderson, John M. & Colin J. Ewen 1987. *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, John M. & Charles Jones 1972. Three theses concerning phonological representations. *Journal of Linguistics* 10: 1–26.

Anderson, Stephen 1985. *Phonology in the Twentieth Century*. Chicago: University of Chicago Press.

Archangeli, Diane 1988. Aspects of underspecification theory. *Phonology* 5: 183–208.

Archangeli, Diane & Douglas Pulleyblank 1989. Yoruba vowel harmony. *Linguistic Inquiry* 20.2: 173–217.

Bach, Emmon & Robert T. Harms 1972. How do languages get crazy rules? *Linguistic Change and Generative Rheory*, edited by Robert Stockwell & Ronald Macaulay, 1–21. Bloomington: Indiana University Press.

Backley, Phillip 1993. Coronal: The undesirable element. *UCL Working Papers in Linguistics* 5: 301–323.

Backley, Phillip 2002. Prosodic markers in melodic representation. *Studies in Languages and Cultures* 16. Faculty of Languages and Cultures, Kyushu University, 125–145.

Backley, Phillip 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Backley, Phillip & Kuniya Nasukawa, 2009. Headship as melodic strength. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 47–77. Berlin and New York: Mouton de Gruyter.

Bellem, Alex 2007. *Towards a Comparative Typology of Emphatics: Across Semitic and into Arabic Dialect Phonology*. Ph.D. dissertation, SOAS.

Bermúdez-Otero, Ricardo 2007. Diachronic phonology. *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 497–518. Cambridge: Cambridge University Press.

Blaho, Sylvia, Patrick Bye & Martin Krämer (eds.) 2007. *Freedom of Analysis?* Berlin: Mouton de Gruyter.

Botma, Bert 2004. *Phonological Aspects of Nasality*. Ph.D. dissertation, University of Amsterdam. Utrecht: LOT Publications.

Botma, Bert 2009. Transparency in nasal harmony and the limits of reductionism. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 79–112. Berlin and New York: Mouton de Gruyter.

Botma, Bert, Nancy C. Kula & Kuniya Nasukawa 2013. Features. *Bloomsbury Companion to Phonology*, edited by Nancy C. Kula, Bert Botma & Kuniya Nasukawa, 33–63. London: Bloomsbury.

Botma, Bert & Norval S. H. Smith 2006. A dependency account of the fortis-lenis contrast in Cama. *Linguistics in the Netherlands 2006*, edited by Jeroen M. van de Weijer & Bettelou Los, 15–27. Amsterdam and Philadelphia: John Benjamins.

Botma, Bert & Norval S. H. Smith 2007. A dependency-based typology of nasalization and voicing phenomena. *Linguistics in the Netherlands 2007*, edited by Bettelou Los & Marjo van Koppen, 36–48. Amsterdam and Philadelphia: John Benjamins.

Bradshaw, Janet 2003. *Consonant-Tone Interaction in Siwati*. Ms., Kyungpook National University.

Bromberger, Sylvain & Morris Halle 1989. Why phonology is different. *Linguistic Inquiry* 20: 51–70.

Charette, Monik 1990. License to govern. *Phonology* 7: 223–253.

Charette, Monik 1991. *Conditions on Phonological Government*. Cambridge: Cambridge University Press.

Charette, Monik & Asli Göksel 1994. Vowel harmony and switching in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 4: 31–52. Also in Kardela, Henryk, Bogdan Szymanek (eds.), *A Festschrift for Edmund Gussmann*, 29–56. Lublin 1996: University Press of the Catholic University of Lublin. WEB.

Charette, Monik & Asli Göksel 1996. Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 6: 1–25. Also in Cyran, Eugeniusz (ed.), *Structure and Interpretation: Studies in Phonology*, 65–88. Lublin 1998: Folium. WEB.

Chomsky, Noam 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam 1973. Conditions on transformations. *A Festschrift for Morris Halle*, edited by Stephen Anderson & Paul Kiparsky, 232–286. New York: Holt, Rinehart & Winston.

Chomsky, Noam & Morris Halle 1968. *The Sound Pattern of English*. New York: Harper and Row.

Clements, George N. 1981. Akan vowel harmony: A non-linear analysis. *Harvard Studies in Phonology* 2: 108–177.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2: 225–252.

Cobb, Margaret 1997. *Conditions on Nuclear Expressions in Phonology*. Ph.D. dissertation, SOAS, University of London.

Cyran, Eugeniusz 1997. *Resonance Elements in Phonology: A Study in Munster Irish*. Lublin: Folium.

Cyran, Eugeniusz 2010. *Complexity Scales and Licensing in Phonology*. Berlin and New York: Mouton de Gruyter.

Cyran, Eugeniusz 2014. *Between Phonology and Phonetics: Polish Voicing*. Berlin: Mouton de Gruyter.

de Lacy, Paul 2002. *The Formal Expression of Markedness*. Ph.D. dissertation, University of Massachusetts.

de Lacy, Paul 2007. Themes in phonology. *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 5–30. Cambridge: Cambridge University Press.

Dresher, Elan B. 2009. *The Contrastive Hierarchy in Phonology*. (Cambridge Studies in Linguistics 121.) Cambridge: Cambridge. University Press.

Dresher, Elan B. & Richard Compton 2011. Palatalization and "strong i" across Inuit dialects. *Canadian Journal of Linguistics* 56: 203–228.

Embick, David & Rolf Noyer 2007. Distributed morphology and the syntax–morphology interface. *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand & Charles Reiss, 289–324. Oxford: Oxford University Press.

Fodor, Jerry 1983. *The Modularity of the Mind*. Cambridge, MA: MIT-Bradford.

Gardner, Howard 1985. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Gussmann, Edmund 2002. *Phonology: Analysis and Theory*. Cambridge: Cambridge University Press.
Gussmann, Edmund 2007. *The Phonology of Polish*. Oxford: Oxford University Press.
Hale, Mark & Charles Reiss 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
Hall, Tracy A. 2007. Segmental features. *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 311–334. Cambridge: Cambridge University Press.
Harris, John 1990. Segmental complexity and phonological government. *Phonology* 7: 255–300.
Harris, John 1994. *English Sound Structure*. Oxford: Blackwell. WEB.
Harris, John & Geoff Lindsey 1990. Phonetic interpretation in Generative Grammar. *UCL Working Papers in Linguistics* 2: 355–369.
Harris, John & Geoff Lindsey 1995. The elements of phonological representation. *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 34–79. Harlow, Essex: Longman. WEB.
Hooper, Joan 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
Huber, Daniel 2003. Velars and empty-headedness in government phonology. *Rivista di Grammatica Generativa* 28: 45–56.
Hulst, Harry van der 1989. Atoms of segmental structure: Components, gestures and dependency. *Phonology* 6: 253–284.
Hulst, Harry van der 1994. Radical CV Phonology: The locational gesture. *UCL Working Papers in Linguistics* 6: 439–477.
Hulst, Harry van der 1995. Radical CV phonology: The categorial gesture. *Frontiers of Phonology: Atoms, Structures, Derivations*, edited by Jacques Durand & Francis Katamba, 80–116. London and New York: Longman.
Hulst, Harry van der 2000. Modularity and modality in phonology. *Phonological Knowledge: Conceptual and Empirical Issues*, edited by Noel Burton-Roberts, Philip Carr & Gerard Docherty, 207–243. Oxford: Oxford University Press.
Itô, Junko & Armin Mester 1986. The phonology of voicing in Japanese. *Linguistic Inquiry* 17: 49–73.
Jakobson, Roman, Gunnar Fant & Morris Halle 1952. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates* (Second edition). Cambridge, MA: MIT Press.
Jakobson, Roman & Morris Halle 1956. *Fundamentals of Language*. The Hague: Mouton.
Jensen, Sean 1994. Is ʔ an element? Towards a non-segmental phonology. *SOAS Working Papers in Linguistics & Phonetics* 4: 71–78.
Kaye, Jonathan 1982. Harmony processes in Vata. *The Structure of Phonological Representations*, edited by Harry van der Hulst & Norval Smith, Vol 2, 385–452. Dordrecht: Foris.
Kaye, Jonathan 1989. *Phonology: A Cognitive View*. Hillsdale: Erlbaum. WEB.
Kaye, Jonathan 1990. What ever happened to dialect B? *Grammar in Progress: GLOW Essays for Henk van Riemsdijk*, edited by Joan Mascaró & Marina Nespor, 259–263. Dordrecht: Foris.
Kaye, Jonathan 1992. On the interaction of theories of Lexical Phonology and theories of phonological phenomena. *Phonologica 1988*, edited by Uli Dressler, Hans Luschützky, Oskar Pfeiffer & John Rennison, 141–155. Cambridge: Cambridge University Press. WEB.
Kaye, Jonathan 1995. Derivations and interfaces. *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 289–332. London & New York: Longman. Also in *SOAS Working Papers in Linguistics and Phonetics* 3, 1993, 90–126. WEB.
Kaye, Jonathan 2000. *A Users' Guide to Government Phonology*. Ms., University of Ulster. WEB.
Kaye, Jonathan 2001. Working with licensing constraints. *Constraints and Preferences*, edited by Katarzyna Dziubalska-Kołaczyk, 251–268. Berlin & New York: Mouton de Gruyter. WEB.
Kaye, Jonathan 2005. GP, I'll have to put your flat feet on the ground. *Organizing Grammar: Studies in Honor of Henk van Riemsdijk*, edited by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster, 283–288. Berlin: Mouton de Gruyter.
Kaye, Jonathan 2008. *Canadian Raising, eh?* Ms.
Kaye, Jonathan & Jean Lowenstamm 1981. Syllable structure and markedness theory. *Theory of Markedness in Generative Grammar*, edited by Adriana Belletti, Luciana Brandi & Luigi Rizzi, 287–315. Pisa: Scuola Normale Superiore. WEB.

Kaye, Jonathan & Jean Lowenstamm 1984. De la syllabicité. *Forme Sonore du Langage*, edited by François Dell, Daniel Hirst & Jean-Roger Vergnaud, 123–159. Paris: Hermann. WEB.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud 1985. The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2: 305–328. WEB.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. *Phonology* 7: 193–231. WEB.

Kiparsky, Paul 1968–1973. How abstract is phonology? Manuscript circulated since 1968 and published 1973. *Three Dimensions of Linguistic Theory*, edited by Osamu Fujimura, 5–56. Tokyo: TEC.

Kiparsky, Paul 1974. On the evaluation measure. *Papers from the Parasession on Natural Phonology*, edited by Anthony Bruck, Robert Allen Fox & Michael W. La Galy, 328–337. Chicago: Chicago Linguistic Society.

Kiparsky, Paul 1982. Lexical morphology and phonology. *Linguistics in the Morning Calm*, edited by In-Seok Yang, 3–91. Seoul: Hanshin.

Kula, Nancy C. 1999. On the representation of NC clusters in Bemba. *Linguistics in the Netherlands 1999*, edited by Renée van Bezooijen & René Kager, 135–148. Amsterdam and Philadelphia: John Benjamins.

Kula, Nancy C. 2002. *The Phonology of Verbal Derivation in Bemba*. Ph.D. dissertation, University of Leiden. Utrecht: LOT Publications.

Kula, Nancy C. 2005. On licensing constraints for consonants. In Nancy C. Kula & Jeroen van de Weijer (eds.) *Proceedings of the Government Phonology Workshop*. *Special issue of Leiden Papers in Linguistics* 2.4: 51–75.

Kula, Nancy C. 2012. On the representation of tone in Element Theory. *Sound, Structure and Sense*, edited by Eugeniusz Cyran, Henryk Kardela & Bogdan Szymanek, 353–367. Lublin: Wydawnictwo KUL.

Kula, Nancy C. & Lutz Marten 1998. Aspects of nasality in Bemba. *SOAS Working Papers in Phonetics and Linguistics* 8: 191–208.

Kula, Nancy C. & Lutz Marten 2000. Constraints and processes: Evidence from Bemba, Herero and Swahili. *SOAS Working Papers in Phonetics and Linguistics* 10: 91–102.

Lowenstamm, Jean 1999. The beginning of the word. *Phonologica 1996: Syllables !?*, edited by John R. Renison & Klaus Kühnammer, 153–166. The Hague: Holland Academic Graphics.

Lowenstamm, Jean & Jonathan Kaye 1986. Compensatory lengthening in Tiberian Hebrew. *Studies in Compensatory Lengthening*, edited by Leo Wetzels & Engin Sezer, 97–132. Dordrecht: Foris.

McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 43: 84–108.

Nasukawa, Kuniya 1995. Nasality and harmony in Gokana. *UCL Working Papers in Linguistics* 7: 511–533.

Nasukawa, Kuniya 1997. Melodic structure in a nasal-voice paradox. *UCL Working Papers in Linguistics* 9: 403–423.

Nasukawa, Kuniya 1998. An integrated approach to nasality and voicing. *Structure and Interpretation: Studies in Phonology*, edited by Eugeniusz Cyran, 205–225. Lublin: Folium.

Nasukawa, Kuniya 2005a. *A Unified Approach to Nasality and Voicing*. Berlin and New York: Mouton de Gruyter.

Nasukawa, Kuniya 2005b. The representation of laryngeal-source contrasts in Japanese. *Voicing in Japanese*, edited by Jeroen van de Weijer, Tetsuo Nishihara & Kensuke Nanjo, 79–99. Berlin and New York: Mouton de Gruyter.

Nasukawa, Kuniya 2010. No consonant-final stems in Japanese verb morphology. *Lingua* 120: 2336–2352.

Nasukawa, Kuniya & Phillip Backley 2005. Dependency relations in element theory: Markedness and complexity. In Nancy C. Kula & Jeroen van de Weijer (eds.), *Proceedings of the Government Phonology Workshop*. *Special issue of Leiden Papers in Linguistics* 2.4: 77–93.

Nasukawa, Kuniya & Phillip Backley 2008. Affrication as a performance device. *Phonological Studies* 11: 35–46.

Newmeyer, Frederick 1986. *Linguistic theory in America*. 2nd edition. New York: Academic Press.
Oostendorp, Marc van 2005. The first person singular in Dutch dialects. *Proceedings of the Thirty-Fifth Annual Meeting of the North East Linguistic Society*, edited by Leah Bateman & Cherlon Ussery, 1–12. Amherst, MA: GLSA.
Paradis, Caroline & Jean François Prunet 1991. *The Special Status of Coronals: Internal and External Evidence*. San Diego: Academic Press.
Pearce, Mary 2009. Kera tone and voicing interaction. *Lingua* 119: 846–864.
Ploch, Stefan 1999. *Nasals on My Mind: The Phonetic and the Cognitive Approach to the Phonology of Nasality*. Ph.D. dissertation, SOAS.
Pöchtrager, Markus 2006. *The Structure of Length*. Ph.D. dissertation, University of Vienna. WEB.
Pöchtrager, Markus & Jonathan Kaye 2013. GP2.0. *SOAS Working Papers in Linguistics* 16: 51–64.
Polgárdi, Krisztina 2009. Trochaic proper government, loose CV, and vowel-zero alternation in Hungarian. *Approaches to Hungarian, Vol.11: Papers from the 2007 New York Conference*, edited by Marcel den Dikken & Robert Vago, 143–165. Amsterdam: Benjamins.
Rennison, John 1990. On the elements of phonological representations: The evidence from vowel systems and vowel processes. *Folia Linguistica* 24: 175–244.
Rennison, John & Friedrich Neubarth 2003. An x-bar theory of government phonology. *Living on the Edge: 28 Papers in Honour of Jonathan Kaye*, edited by Stefan Ploch, 95–130. Berlin and New York: Mouton de Gruyter.
Ritter, Nancy 1997. Headedness as a means of encoding stricture. *Phonology in Progress – Progress in Phonology*, edited by Geert Booij & Jeroen van de Weijer, 333–365. Leiden: Holland Academic Graphics.
Sagey, Elizabeth 1986. *The Representation of Features and Relations in Non-Linear Phonology*. Ph.D. dissertation, MIT.
Schane, Sanford 1984. The fundamentals of particle phonology. *Phonology Yearbook* 1: 129–155.
Scheer, Tobias 1996. *Une théorie de l'interaction directe entre consonnes*. Doctoral dissertation, Université Paris 7.
Scheer, Tobias 1999. A theory of consonantal interaction. *Folia Linguistica* 32: 201–237.
Scheer, Tobias 2004. *A Lateral Theory of Phonology: Vol.1: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.
Scheer, Tobias 2010. What OT is, and what it is not: Review of the Cambridge handbook of phonology, ed. by Paul de Lacy. *Journal of Linguistics* 46: 193–218. WEB.
Scheer, Tobias 2011. Aspects of the development of generative phonology. *The Continuum Companion to Phonology*, edited by Bert Botma, Nancy Kula & Kuniya Nasukawa, 397–446. New York: Continuum. WEB.
Scheer, Tobias 2015. How diachronic is synchronic grammar? Crazy rules, regularity and naturalness. *The Handbook of Historical Phonology*, edited by Patrick Honeybone & Joseph C. Salmons, 313–336. Oxford: Oxford University Press.
Scheer, Tobias 2016. Melody-free syntax and phonologically conditioned allomorphy. *Morphology* 26: 341–378.
Segal, Gabriel 1996. The modularity of theory of mind. *Theories of Theories of Mind*, edited by Peter Carruthers & Peter K. Smith, 141–157. Cambridge: Cambridge University Press.
Smith, Jennifer 2011. Category-specific effects. *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin Ewen, Beth Hume & Keren Rice, 2439–2463. Malden, MA: Wiley-Blackwell.
Smith, Norval H. S. 1988. Consonant place features. *Features, Segmental Structure and Harmony Processes, Part I*, edited by Harry van der Hulst & Norval S. H. Smith, 209–236. Dordrecht: Foris.
Stampe, David 1972. *How I Spent my Summer Vacation*. Ph.D. dissertation, University of Chicago. Reproduced by the Indiana Linguistics Club in 1979.
Steriade, Donca 2007. Contrast. *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 139–157. Cambridge: Cambridge University Press.

Szigetvári, Péter 1994. Coronality, Velarity and why they are special. *The Even Yearbook* 1: 185–224.
Tranel, Bernard 1981. *Concreteness in Generative Phonology: Evidence from French*. Berkeley: University of California Press.
Turing, Alan 1936–37. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, series 2, 42: 230–265.
Živanović, Sašo & Markus A. Pöchtrager 2010. GP 2, and Putonghua too. *Acta Linguistica Hungarica* 57.4: 357–380.

10

Syllable structure in Government Phonology

Tobias Scheer and Eugeniusz Cyran

10.1 The core of GP: lateral relations

10.1.1 Lateralization of structure and causality in two steps

Traditionally, syllable structure is encoded by arboreal structure, i.e. the syllabic tree with the canonical constituent structure [Onset [Nucleus Coda]$_{\text{Rhyme}}$]$_{\sigma}$. The core of the research programme of Government Phonology (GP) is to show that the syllabic position of a segment is not defined by a constituent to which it belongs (and whose status is itself defined by the arboreal relations that it entertains with other constituents), but by lateral relations that hold among constituents.

This is where the name of the theory comes from, which is an indication (also from hindsight) that lateral relations are at the heart of GP: government is one such lateral relation, licensing is another. Government was introduced by Kaye et al. (1990) with explicit reference to government in the then current GB syntax where the well-formedness of movement was defined as a (lateral) relationship between the base (the trace, an empty constituent) and the target position of the displaced item called government. Syntactic government was sensitive to intervening structure (barriers).

Syllable structure and syllabic causality (i.e. the reason why a segment reacts on syllabic pressure) were lateralized in two steps, with an intermediate stage. Standard GP introduced lateral relations and shifted a certain amount of labour from the syllable tree to this new device. As a result, arboreal structure was severely impoverished, but remained in a depleted guise. This led to a hybrid model where on many occasions the same labour was done twice: by lateral and by arboreal structure. Conceptually, this kind of situation is undesirable, and this was made explicit early on by Takahashi's (1993) paper *A Farewell to Constituency* where the author shows that there is no need for arboreal structure when lateral relations are in place because the former can be read off the latter. Or, put differently, arboreal structure is a notational variant of lateral relations, which are primary. Hence the syllabic tree is a mere consequence of lateral relations and as such enjoys no theoretical status (but may help bridge the gap graphically between the old and the new for linguists who are familiar with trees). A trademark of GP, empty nuclei, are a direct consequence of this first step in the lateralization of structure and causality.

John Harris' (1997) Licensing Inheritance modifies the original hybrid model by adding more lateral relations, thereby completing the lateralization of causality – but without touching arboreal structure. This was a further step on the lateral trajectory that GP has initiated, but in a sense made the arboreal-lateral tension even more acute (see Scheer 2004: §§172, 186).

The last piece to be lateralized was thus the remaining arboreal structure itself. If it is understood that a theory cannot afford to have the same labour done twice, by (primary) lateral relations and (secondary) arboreal structure, the latter has to go if the lateral project is worthwhile at all. This is what Lowenstamm (1996) sets out to do: arboreal syllable structure is done away with altogether (constituency reduces to a strict sequence of non-branching onsets and non-branching nuclei), and lateral relations alone define syllabic positions. This is the lateral project described in Scheer (2004) (especially §165, where the evolution from Standard GP to CVCV is traced back in detail; on this transition see also Scheer 2012a), which is known as *CVCV* or *Strict CV* (see section 10.2 below).

More recently, syllable structure along the lines of Standard GP is used e.g. by Charette (2008). It also underlies GP 2.0 (Pöchtrager 2006; Pöchtrager & Kaye 2013), even though this approach focuses on melodic structure (aiming to further reduce the set of melodic primes by expressing their contribution in terms of arboreal structure; see Chapter 9, section 9.2.1).

10.1.2 A pre-theoretical fact: lateral relations encode the workings of syllabification

Syllable structure is a function of two and only two factors: the order of segments in the linear string and their sonority with respect to their neighbours. This is an uncontroversial fact which all phonological theories implement in one way or another (whatever formal guise 'sonority' might take). The two linear strings VTRV and VRTV (T is shorthand for any obstruent, R for any sonorant), for example, will end up syllabified as V.TRV and VR.TV because TR makes a good branching ONSET while RT does not. The well-formedness of branching onsets, in turn, is defined in terms of sonority: only rising sonority profiles qualify. What that means is that the decision on syllabification, i.e. the labour of syllabification algorithms in traditional theories, relies on the evaluation of the relationship that is entered into by adjacent segments, here by two adjacent consonants: the relative sonority of TR and RT produces different outputs. The decision is thus *lateral* in kind: given linearity, it is based exclusively on the comparison of the sonority values of the two items at hand.[1]

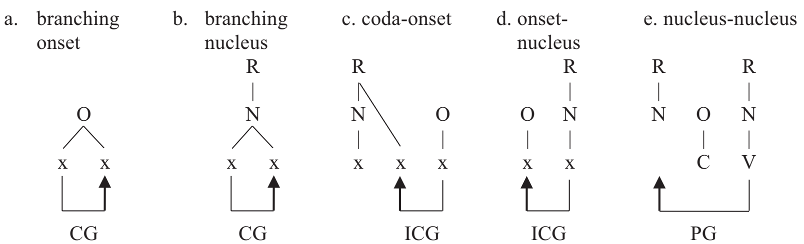
Rather than talking about the effect of syllabification (constituent structure), GP elevates to theoretical relevance its origin: the lateral relationship between two segments. Or, in other words, GP encodes the lateral cause of syllabification, rather than its vertical (arboreal) effect: it is the former that matters and that phonological theory should manipulate. GP thus offers a direct snapshot of the driving mechanism, rather than a picture of its indirect consequences (see Scheer 2004: §11). In sum, everybody is doing the same lateral computation when doing syllable structure, but only GP is making this explicit.

The relationship among two adjacent consonants regarding their sonority is encoded in Standard GP by two lateral forces: Constituent Government (CG) and Interconstituent Government (ICG). Sonority is encoded in terms of segmental complexity, i.e. the number of primes that a segment is made of (see Harris 1990; Chapter 9, section 2.2). The general rule, then, is that more complex segments (which are stronger because they bear more primes) are more important than less complex items when their relationship as neighbours is calculated: being more important means heading the cluster, and the asymmetric, hierarchical relationship between a head and a complement is called government. Since obstruents are found to be systematically more complex than sonorants (see Chapter 9, section 2.2), they are the head of clusters where they occur with sonorants, to the effect that TR clusters are left-headed (CG), while RT clusters are right-headed (ICG).

Note that the name of the two kinds of government refers to the arboreal output of the computation (TR ends up as a branching Onset, RT as a coda–onset cluster), rather than to their lateral properties, which are opposite in terms of direction: CG is progressive, while ICG is regressive. The directionality of government determines constituency more generally in Standard GP since complex nuclei (long vowels or heavy diphthongs) are also left-headed (and hence again within a constituent, where government is left-to-right), while government among different nuclei is right-to-left. Internuclear government has a special name, Proper Government, and is at the origin of vowel-zero alternations, as we will see shortly. For the time being, table (1) below recapitulates the different types of government that are recognized in Standard GP (a more detailed short guide to 1990 GP is available in Scheer 2004: §623).

(1) Government in Standard Government Phonology

Constituent Government (CG) relates two members of the same constituent	Interconstituent Government (ICG) relates two segments that pertain to different (consonantal) constituents	Proper Government relates two distinct nuclei



10.1.3 *Arboreal structure depleted: no ternary constituents, no syllable node, no coda constituent, no word-final codas*

The lateral perspective on syllable structure has a number of consequences, which all concur to reduce arboreal structure (with respect to the traditional standard). For example, ternary constituents are ruled out. This follows from the fact that (Proper Government set aside, on which see section 10.1.5) relative sonority and hence lateral relations are only ever computed among two adjacent segments – an observation called *strict locality* in Standard GP (recall that this embodies the basic fact that syllable structure is a computation of the relative sonority of adjacent segments). For a putative ternary constituent [x_1 x_2 x_3], strict locality requires x_2 to be the head (otherwise two non-adjacent items would be related). However, x_2 could not govern x_1 because we know that within a constituent, government is always progressive (something that is called *strict directionality* in Standard GP). Hence the impossibility for the system to derive well-formed ternary constituents, which is known as the *Binary Theorem* (Kaye et al. 1990: 199f.; Kaye 1990: 306f.): all syllabic constituents are maximally binary.

Another consequence of the lateral system is that there cannot be any constituent dominating the Onset and the rhyme: there is no syllable node (Kaye et al. 1990: 200f.;

Brockhaus 1995). Were the Onset and the rhyme sisters within a constituent, the former would have to govern the latter since CG is progressive. For one thing, this sounds awfully odd to a phonologist's ear since the fundamental asymmetry between consonants and vowels is the other way around: the centre of gravity of the syllable is the vowel; consonants are satellites. But setting aside intuition, this scenario would also violate the Binary Theorem since only segments are governors and governees. In a syllable with a branching Onset and a branching nucleus, it is unclear which one of the two consonants would govern which one of the two vocalic slots, and strict locality would be violated anyway.

As a result, Standard GP thus claims that there is no syllable node. Instead there are onset–rhyme pairs. This situation has fed an urban myth saying that GP has no syllable or is against the syllable. In the 1990s, this myth was wide-spread among phonologists (and still persists in some quarters), showing that GP texts were not directly perceived in the community. Having no syllable node does not mean being uninterested in syllable structure, or being against this very concept. There are various theories of sound structure, namely found in phonetic quarters or in the trend towards a 'phonetically grounded phonology' that was popular in the early 2000s (Steriade 1999; Hayes et al. 2004), which truly propose that there is no such thing as syllable structure (since it may be predicted from phonetic cues). GP is quite the opposite of that (recall from Chapter 9.1.1 that for Kaye phonology is a system that computes cognitive units).

A further consequence of the lateral system is to demote the coda to a non-branching existence, and ultimately to a non-constituent (Kaye et al. 1990: 201f.). The bell-curve describing the syllable that was identified at least since Sievers (1885) has a rising sonority slope on the left side of the nucleus, but a falling profile to its right. That is, candidates for branching codas show a falling sonority slope (*carp*, *salt* etc.). Since obstruents are governors and sonorants governees, the only governing relation that could hold within the cluster rp of *carp* is one where the *p* governs the *r*. This, however, violates strict directionality: government within a constituent, hence within a putative branching coda, is head-initial. Another consequence of the lateral system is that the coda would be the only constituent that never governs anything (onsets govern preceding codas, nuclei govern their onsets). It is therefore denied the status of a constituent (Kaye et al. 1990: 201f.), and its skeletal slot is directly attached to the rhyme. Instead of *coda*, then, it is referred to as *postnuclear rhymal complement* or *rhymal adjunct* (for expository reasons, the word *coda* continues to be used below, though).

Finally, the status of coda consonants as being governed by the following Onset prompts an obvious question: what about word-final codas? What would be their governor? It was mentioned that codas are only ever dependents, i.e. they never govern anything. From this Kaye (1990) draws the conclusion that codas cannot exist without being governed by a following Onset. Which means that there are no word-final codas at all, in any language, since they will never be able to be governed. If word-final consonants cannot be codas, there is only one alternative: they are onsets. Which means in turn that there must be a nucleus to their right: the existence of an Onset implies the existence of its nucleus. This nucleus, then, must be empty since the word is consonant-final on the surface. A word such as *cat* thus ends in a final empty nucleus (FEN): /katø/.[2] In this setting, the parametric decision for a language to allow for word-final consonants or not (English says yes, Italian says no, at least in its native vocabulary) is a decision about FEN, rather than about final consonants: languages that have final consonants allow final nuclei to be empty, while languages where all words are vowel-final require that final nuclei be phonetically expressed (Kaye 1990: 323ff.).

Up to this point, the purpose of lateral forces (except for Proper Government, on which more below) was to encode sonority relations among adjacent segments. There is no obligation, however, for a consonant to enter in a sonority relationship with a neighbour: onsets may of course be non-branching and intervocalic. Hence Kaye's conclusion has a new quality: he derives the *obligation* for codas to be governed, i.e. to be unable to exist on their own, from their exclusive status as governees. This obligation is called *Coda Licensing* and introduces a new category of lateral forces, or at least a new word for a subset of them: *licensing*. What is meant is that licensing describes a situation where the target cannot exist in absence of the lateral force in question. But still, Coda Licensing is supposed to be a consequence, or an effect, of ICG: it is the latter that has the inherent virtue to rule over the existence of codas, and this inherent virtue is then called licensing. The relationship between government and licensing thus appears to be one of inclusion: licensing is a virtue, or an effect, of government.

An extension of the new licensing tool is *Government Licensing*, introduced by Charette (1990). Like Coda Licensing, it describes a well-formedness condition on the existence of dependent skeletal slots (these were first consonantal, but the concept was extended to vocalic slots by Yoshida 1993). The idea is that governors do not enjoy their governing ability per se: they need external support in order to be able to dominate their complement. Hence simplex, but not branching onsets, can be followed by an empty nucleus: only contentful nuclei qualify as government-licensors. For the same reason, a simplex Onset that is called to govern a preceding coda must not be followed by an empty nucleus.

In practice, Government Licensing was developed through the analysis of Quebec French where the realization of schwa is optional unless it is preceded by a consonant cluster, in which case its presence is mandatory. The preceding cluster may either be a coda–onset sequence as in *forteresse* [fɔrtərɛs], *[fɔrtrɛs] 'fortress', or a branching Onset as in *autrement* [otrəmã], *[otrmã] 'otherwise'. In both cases, schwa cannot be dropped (in this variety of French) since the head of the preceding consonantal governing domain (the t in both cases) needs to be licensed in order to be able to govern its complement.

In more recent work, the idea of Government Licensing is further developed into licensing strength scales (Cyran 2010).

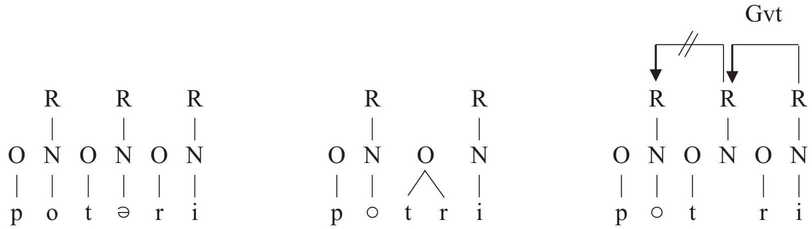
10.1.4 Vowel-zero alternations, empty nuclei and resyllabification

Empty nuclei are a trademark of GP. Even though they were sporadically used outside of and prior to this theory (Anderson 1982; Angoujard 1982; Spencer 1986), only GP has given them a theoretical status with stable diagnostics and cross-linguistic properties. While word-final empty nuclei as described in the previous section have spread into the phonological landscape and today are broadly assumed (e.g. Kiparsky 1991; Dell 1995; Burzio 1994; Oostendorp 2005), internal empty nuclei were outcasts in the 1990s and today are still more or less confined to GP quarters.

Recall from Chapter 9.1.1 that in GP what you get is *not* what you see. Linguists in general and phonologists in particular teach that linguistic structure cannot be read off the surface but needs to be discovered through analysis. In phonemic analysis, for example, what you get may be quite different from what you see since e.g. two phonetic items (segments) may turn out to be just one phonological unit (a phoneme). Strangely enough, though, when it comes to syllable structure the same phonologists may well teach that what you get is only ever what you see: there are exactly as many phonological units (x-slots) as there are phonetic items (segments). Almost all of them will agree that there are empty onsets; some will also admit word-final empty nuclei, but rarely will phonologists provide for internal empty nuclei. Empty constituents, nuclei just as much as onsets, though, are predicted to exist by basic autosegmental workings, whose central insight is that the relationship between syllabic constituents and segments is not one-to-one (Scheer 2015: §93ff.): there must be segments that are unassociated (floating segments), and there must be constituents that have no segmental content (empty onsets and nuclei).

Beyond these principled considerations, internal empty nuclei are the analytic alternative to resyllabification when a vowel alternates with zero. A French word like *poterie* 'pottery' may be pronounced either with (*pot*[ə]*rie*) or without (*pot'rie*) a schwa in the penult syllable. The form with realized schwa appears under (2a), while (2b) and (2c) show the two analyses of the schwaless pronunciation with and without resyllabification.

(2) a. schwa realized — b. schwa absent, resyllabification — c. schwa absent, no resyllabification



In this particular (optional) vowel-zero alternation found in French, we know for sure that there was no resyllabification of the two independent onsets *t* and *r* into a branching Onset. This is because in relevant (southern) varieties of French, +ATR and -ATR versions of mid vowels are in complementary distribution (Durand 1990: 24ff.): the former occur in open (*f[o]lie* 'madness', including before TRs: *se v*[o]*trer* 'to lounge'), while the latter are observed in closed syllables (*div*[ɔ]*rcer* 'to divorce'). The resyllabified structure under (2b) should thus produce an [o] – but what is really pronounced is [ɔ]: *p*[ɔ]*t'rie*, **p*[o]*t'rie*. Hence there was no resyllabification (Scheer 2015: §112f.). The -ATR pronunciation is thus consistent with (2c) where the melody of the schwa was deleted, but not its constituent.[3]

10.1.5 Proper Government and structure preservation

This further depletes arboreal structure: what used to be a branching Onset now identifies as two independent onsets that enclose an empty nucleus. The analysis also establishes a *lateral causality*: the vowel is -ATR because it is followed by an empty nucleus (contentful nuclei produce +ATR vowels: *f*[o]*lie*). Hence the [ɔ] of *p*[ɔ]*t'rie* /potøri/ occurs in a 'closed syllable' as much as the [ɔ] in *div*[ɔ]*rcer* /divorøse/ 'to divorce', that is in fact before a governed empty nucleus. The absence of lateral communication between the empty nucleus and the mid vowel is indicated by a barred arrow under (2c). The arrow is barred because (governed) empty nuclei (as opposed to contentful nuclei) cannot be the source of lateral relations.

The other lateral relation depicted under (2c), i.e. the one between the final vowel and the nucleus that hosts the vowel-zero alternation, indicates that schwa is deleted under government. In Standard GP, internuclear government was called Proper Government in order to be distinct from the other types of (interconsonantal) government mentioned. We will see in section 10.1.8 that the multiplication of distinct lateral forces was also a reason to move on to CVCV, where the lateral zoo boils down to government and licensing (or even, in Cyran's model, to licensing alone; see section 10.2.1). But at this point two things are noteworthy: (i) Proper Government is the only lateral relation in Standard GP which is not strictly local: it relates two skeletal slots that are not adjacent; and (ii) Proper Government is the only type of government that does not describe the sonority relationship of segments: it is precisely independent of the sonority slope of consonants that surround the alternation site.

But let us return to the French case. Note that word-final 'closed syllables' are also covered: *folle* [fɔl] 'mad, fem.' displays a -ATR vowel before a word-final consonant. Recall that according to GP this consonant is the ONSET of an empty nucleus. Hence the overall description of the two positions where [ɔ] occurs is unified (i.e. non-disjunctive): [ɔ] is observed iff the following nucleus is empty.[4] This is the essence of what Kaye (1990) establishes on the grounds of vowel-zero alternations in Yawelmani and Turkish. Note that he did not take the step to identify __C{C,#} for what it classically is, though: a closed syllable. This is because Standard GP had remaining arboreal structure, including codas, which literally (and in the classical sense) closed the syllable. A consequence of this stance is that Standard GP had two distinct identities for vowels in closed syllables: they could occur 'before an empty nucleus' and 'before a coda'. It is this kind of redundancy that will lead CVCV to eliminate all remnants of arboreal structure in the 1990s, and hence to simply say that a vowel stands in a closed syllable iff the following nucleus is empty (see section 10.2.2 below).

Finally, stepping back from the singular French pattern, there are certainly cases where the analyst is not as lucky as in this language, which allows him to control the syllabic status of the consonant preceding the alternation site by looking at the preceding vowel. A position according to which resyllabification occurs in absence of compelling evidence to the contrary (of the French kind) is thus empirically sound. Rather than menially following the empirical record, though, GP takes cases of the French kind as indicative of the true nature of vowel-zero alternations which may not be revealed in all languages. That is, vowel-zero alternations are purely melodic: they *never* modify syllable structure. When the zero occurs in an alternation site, its syllabic identity is always an empty nucleus.

This means that syllable structure is constant and remains unmodified under phonological processing. In syntax, this idea is known as *structure preservation*: it is not because you don't see an item that its structure (constituent) is not there. This was the basic insight that paved the way for a central tenet of syntactic theory: the well-formedness of a structure depends on the relationship between the place where an item is pronounced (after movement) and the place where it is interpreted (or base-generated, i.e. its trace). The passive transformation once deleted the original constituent of *Peter* in *Peter*$_i$ *is loved t*$_i$ *by Mary* after having moved *Peter*. Structure preservation suspends deletion of unpronounced items (Emonds 1976) and thus produces the same effect as the GP ban on resyllabification. Kaye & Lowenstamm (1984: 125) explicitly refer to this syntactic precedent.

The prohibition against modifying syllable structure in the course of phonological computation is general and absolute in GP. Again in explicit reference to syntax, this was encoded in the *Projection Principle* (Kaye et al. 1990: 224f.). Another typical locus of resyllabification is the right edge of morphemes: on the traditional analysis, C_2 of a C_1VC_2 root will be a coda when pronounced as such (e.g. the k of *luck*), but becomes an ONSET if a vowel-initial suffix is added (as in *luck-y*): $C_1V.C_2V$. This is ruled out in GP: a melodic object that is 'born' in a coda cannot surface in an ONSET (and *vice versa*). Since morpheme-final consonants are always onsets of empty nuclei, the root *luck-* will be /lʌkø/ and the k an ONSET no matter what happens during phonological computation. The suffix *-y* thus enters the FEN.

10.1.6 Syllable structure is recorded in the lexicon

This leads us to another salient property of GP that sets this theory apart from others: the fact that syllable structure is recorded in the lexicon (rather than built by a syllabification algorithm upon phonological computation). This perspective is first mentioned in Kaye & Lowenstamm (1984: 125): "all syllabic constraints, formal and substantive, are defined at the lexical level" (translation mine). If the Projection Principle rules out the modification of syllable structure (i.e. resyllabification), an obvious question is to which derivational stage this ban applies. In systems where computation is based on extrinsically ordered rules, syllabification is a regular piece of phonological computation that may be interleaved with regular segmental rules (say, palatalization). Hence there is no single reference point in the derivation where syllable structure is exhaustively established and then other rules apply. In GP, this reference point is the lexicon, i.e. the pieces that are stored in long-term memory. That is, all units which are stored in the lexicon are fully syllabified and syllabic affiliation may then not be altered during phonological computation. The fact that syllabic items are fully syllabified then extends the Onset status of word-final consonants (that was established by Coda Licensing) to morpheme-final consonants: whether the final *-k* of the root *luck-* ends up being word-final, followed by a consonant or followed by a vowel on the surface is irrelevant – it is an Onset in the lexicon and will have this syllabic affiliation all through.

A common misunderstanding in this context is the conclusion that GP does not have any syllabification algorithm. Of course it does: the soundwave does not come with syllabic marking. Hence infants when acquiring words and adults when learning new lexical items need to syllabify whatever comes their way. In other words, syllabification occurs upon lexicalization, and relevant decisions are made then. These may prove wrong and in this case will be corrected: recall from Chapter 9.1.1 that the syllabic status of the input [wat] in French is ambiguous: the *w* may belong to an Onset or a (diphthongal) nucleus, and this produces two distinct lexical items, *watt* 'watt' and *ouate* 'cotton wool' (the former prohibiting, the latter requiring elision). When a child (or an adult) hears the string [nœf wat] meaning 'nine cotton wools' (*neuf ouates*) and has never come across the word *ouate* before, s/he needs to decide whether the *w* will be lexicalized as an Onset or as the first member of a (nuclear) diphthong. In case the former choice is made, this lexicalization will have to be corrected upon exposure to *l'ouate* 'the cotton wool' where the elision of the definite article *la* shows that the initial Onset of *ouate* is empty.

10.1.7 The growing inventory of empty nuclei

Finally, let us discuss the inventory of empty nuclei that are recognized in Standard GP. Two locations have already been discussed: after morpheme-final consonants and in places where vowels alternate with zero in case the zero surfaces. To complete the picture, there are three more cases to be mentioned. In every instance, empty nuclei are the fall-back analysis when no other syllabic interpretation of a cluster is possible.

In English, word-internal *tl*, *dl* as in *atlas*, *Atlantic*, *butler*, *medley*, *bedlam* etc. are so-called bogus clusters (Gussmann 2002: 75). Their rising sonority slope qualifies them as branching onsets, but they do not occur word-initially (*#tl, *#dl is a pervasive distributional gap in many languages, including Romance and Germanic), a fact that counter-indicates Onset status. Also, the *t* is not aspirated in pre-tonic position where voiceless stops always show aspiration when occurring in an Onset: the *t* of *p*^{h}*ólitics* is unaspirated, but before a tonic vowel in *p*^{h}*olit*^{h}*ícian* it aspirates, and this is also true when voiceless stops occur in a branching

ONSET (*Pat*^{h}*rícia*). In pre-tonic coda position, however, they remain unaspirated (**fac*^{h}*tórial*, **doc*^{h}*tórial*), and so they do in **At*^{h}*lántic*. This indicates coda status for the *t*, which however is ruled out by the sonority slope of *tl*, *dl*. Therefore, if these clusters cannot be either branching onsets or coda-onset clusters, they must be two independent onsets (/atølas/ etc.).

A second case in point involves word-initial s+C sequences. Kaye (1992) shows that these notoriously misbehaving clusters consistently do not behave like branching onsets across languages (see also Harris 1994: 57ff. for English). In Classical Greek reduplication, for example, singleton consonants (*ly-oo – le-ly-ka* 'to lose, present, perfect') including *s* (*satt-oo – se-sag-mai* 'to equip, id.') reduplicate, as does the obstruent of regular branching onsets (*graf-oo – ge-graf-a* 'to write, id.'). This is not the case with *s* when occurring in an s+C cluster: *spoudaz-oo – e-spouda-ka* 'to haste, id.', *stere-oo – e-steree-ka* 'to deprive, id.'. Kaye (1992) concludes that s+C clusters do not instantiate a branching ONSET. This is also indicated by their sonority slope, which notoriously violates sonority sequencing in languages that otherwise restrict initial clusters to #TR. Solutions that have been favoured in the literature place the *s* outside of the ONSET either in a specific constituent (appendix, e.g. Kenstowicz 1994: 260f.) or as a floating item which after syllabification is directly attached to some higher constituent such as the syllable or the prosodic word (Rubach 1999: 292ff.). A status as a contour segment (parallel to affricates) is also sometimes argued for (Steriade 1982: 346ff.). All of these solutions (except the latter, but which does not work for the Greek pattern because it predicts that the entire contour segment will be reduplicated) recur to structures that do not belong to the general syllabic theory used, i.e. which are created only to put up the annoying s+Cs. The alternative is to stick to one's theory, i.e. not to propose a specific patch every time there is a leak. Taking (their) theory seriously is the option that was taken by 19th-century neogrammarians: Sievers (1901: §534) concludes that #s+C clusters instantiate two distinct syllables, hence the leftmost being made of the #s and an empty nucleus. This is also Kaye's (1992) conclusion: the #s is a coda, and the preceding nucleus and ONSET are empty. The fact that Greek reduplicates nothing with s+Cs then follows: in fact it faithfully reduplicates the content of the first ONSET of the word, which happens to be empty. With hindsight, Kaye's (1992) analysis placing an empty nucleus at the left edge of the word is a forerunner of Lowenstamm's (1999) initial CV (see section 10.2.3.5 below).

Finally, Gussmann & Kaye (1993: 448ff.) have devised a mechanism, Interonset Government, in order to account for cases where two empty nuclei occur in a row – a situation that cannot be managed by the tools of the theory as it stands (since Proper Government is not recursive, i.e. vowels can only govern one nucleus at a time). The initial cluster of Polish *mgł-a* [mgw-a] 'mist Nsg' and *pchł-a* [pxw-a] 'flea Nsg' does not qualify as a branching ONSET. It must therefore represent two independent onsets. The following cluster, *gł* [gw] and *chł* [xw], does make a good branching ONSET, but when looking at the genitive plural *mgieł* [mgiɛw], *pcheł* [pxɛw] it turns out that it encloses a vowel-zero alternation. This means that the nominative singular hosts two empty nuclei in a row: /møgøł-a/, /pøchøł-a/. In this situation, Gussmann & Kaye propose that the leftmost empty nucleus is governed by the (suffixal) vowel, while the other empty nucleus is silenced by a governing relationship between the two onsets: since obstruents are governors and sonorants governees, *g* and *ch* [x] govern *ł* [w]. This produces a well-formed structure where both empty nuclei have a reason to remain unpronounced, though at the expense of watering down the analysis of empty nuclei, which were supposed to be a consequence of an internuclear relation (Proper Government). Indeed, the cases at hand would be the only instance of empty nuclei that are due to an interconsonantal relation. Also, Interonset Government violates the directionality of government, which is progressive within a constituent but regressive among constituents (see section 10.1.2).

10.1.8 Increasing labour for the ECP

Summing up the development described, the workings of Standard GP mechanically produce more and more empty nuclei as the languages (and language families) analyzed increase (the Polish case), and as they are applied to notorious problems such as bogus and s+C clusters. Every time a new type of empty nucleus joins in, arboreal syllable structure is washed out a little more. And the growing empty nucleus inventory requires more and more contorted manoeuvres in order to keep everything under one roof: the problematic consequences of Interonset Government were mentioned, and the fact that the vowel following initial s+C clusters needs to govern the empty nucleus preceding the *s* led Kaye (1992) to talk about Magic Government: Proper Government was supposed to be unable to apply over another governing domain such as branching onsets or coda–onset clusters. If word-initial s+C clusters are coda-Onset clusters, though, they are preceded by an empty nucleus whose only reason to remain silent is to be governed by the first expressed nucleus of the word: in /$ø_1ø_2$s.top/ where $ø_2$ is the nucleus of the coda *s* (and $ø_1$ its Onset), $ø_2$ can only be silent because it is governed by the *o*. Since government over a coda–onset cluster is impossible, the solution that fits the empirical picture (*s* in #sC clusters is a coda) is incompatible with the theory. Therefore Kaye calls this government that should not exist magic – not to say that this is the solution and that there is some kind of magic in phonology, but to indicate that this is an unsolved conundrum.

The attempt to manage all these different empty nuclei that occur in different situations and for different reasons was then to say empty nuclei can only occur when there is a good reason for them to remain unpronounced. The idea to place the existence of empty nuclei under grammatical control is yet another borrowing from then current GB syntax: the proliferation of empty categories needs to be marshalled (otherwise they appear everywhere). Hence the Empty Category Principle (ECP) in syntax, whose phonological version in Standard GP simply lists the devices that are able to silence an empty nucleus: (i) word-final empty nuclei are licit upon a parametric decision (see section 10.2.3), and so are empty nuclei that (ii) are properly governed (vowel-zero alternations), (iii) are enclosed within an interonset governing domain and (iv) precede initial s+C clusters (Magic Government). Kaye (1995: 295, 303) coins another lateral relation that the four cases mentioned are supposed to be instantiations of: p-Licensing (where p stands for prosodic). Given this meta-lateral relation, then, empty nuclei are well-formed iff they are p-licensed.

Note that this time various forms of government are a form of (p-)Licensing, while Coda Licensing as described in section 10.2.3 appeared to be a form of government. The unclear relationship between the two major lateral forces, which ought to be distinct but turn out to be forms of one another, could only be managed by a number of patches in an amorphous list. This situation indicated that something is wrong with the theory itself and hence played an important role in the transition to CVCV (see Scheer 2004: §186; section 10.2.3 below).

10.2 Strict CV

10.2.1 Introduction

Strict CV is the final step in the process of lateralization of structure and causality. The elimination of the arboreal aspect of representation and recognition of CV as the only possible syllabic type is now common to a number of variants of Strict CV which differ mainly in how lateral relations between these units are arranged. In this section we look at some current issues connected with the Strict CV model (Lowenstamm 1996, 1999), as developed later in, for example, Scheer (2004); Ségéral & Scheer (2001); Scheer & Ségéral (2008a, 2008b). As a starting point we take the latest version of this model, whose hallmark is Coda Mirror 2 (CM2) (Scheer and Ziková 2010; Scheer 2012a, 2012b). We concentrate on the empirical predictions of this model, its strong and weak points, refer to other 'dialects' of Strict CV, and finally offer some thoughts on a potential further development.

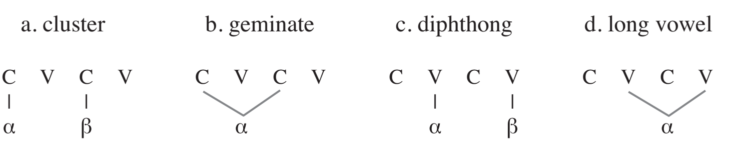
Work in Strict CV includes Lowenstamm (2000a, 2000b, 2003a, 2003b); Bendjaballah (1999, 2001); Bendjaballah & Haiden (2008); Faust (2014, 2015); Passino (2009, 2013); Lahrouchi (2003, 2008); Lahrouchi & Ségéral (2009); Zdziebko (2015); Fortuna (2015, 2016); Ulfsbjorninn (2014); Rizzolo (2007); Barillot & Ségéral (2005); Caratini (2009); and Kula & Marten (2009). The aforementioned dialects of Strict CV include the following. Szigetvári's VC approach holds that the basic building blocks are VC (rather than CV units) and hence the string starts with a V and ends in a C (Szigetvári 1999, 2001, 2007, 2008). Carvalho's Contour CV model defines major classes of segments in terms of a one-to-many association of melodic items to consecutive C- and V-slots (Carvalho 2002, 2004, 2008). Rowicka (1999) and Polgárdi (1999) implement lateral relations that are left-headed (instead of being right-headed). Polgárdi (2003, 2008, 2009) defines a Loose CV environment where the skeleton is strictly CV but ends in a C (there is no FEN). Finally, Cyran (2010) derives the lateral relation network in Strict CV from complexity scales in his Complexity Scales and Licensing model (CSL) where Licensing is the only lateral force (there is no government).

Complexity scales are discussed in section 10.2.4 below, but it is impossible in this chapter to do justice to the wealth of all works and dialects of Strict CV mentioned above.

10.2.2 What is strict CV?

Strict CV itself grew out of a general GP programme (Kaye et al. 1985, 1990; Kaye 1990; Charette 1990; Harris, 1994). Lowenstamm (1996) put forward a proposal that CV is the universal syllable structure, and that any surface departure from this pattern in fact has to be reanalyzed as a phonological sequence of such light syllables. Consequently, consonant clusters, geminates, diphthongs and long vowels must have the following representations. Note that all these configurations enclose an empty position, either C or V.

(3)



Basing his arguments on as diverse phenomena as the template structure of Classical Arabic, Compensatory Lengthening in Tiberian Hebrew or Danish stød, Lowenstamm demonstrates that a number of important generalizations are missed if the above representations are not assumed to be universal. Lowenstamm's paper marks the birth of CVCV, which in a way was a logical consequence of the development of GP. In fact, some Standard GP analyses dealing with, for example, complex consonantal clusters have produced results which were close to the CVCV paradigm, e.g. Gussmann & Kaye (1993) and Cyran & Gussmann (1999) for Polish. The problem, however, was with the proliferation of sometimes contradictory lateral forces to deal with the increased number of empty nuclei (cf. the discussion in section 10.1.8 above).

The simplicity and elegance of the CV assumption is one thing. However, this model dismantles the traditional prosodic structure relating to the syllable and syllabic constituents, which have been referred to as domains of phonotactic interaction between their members, for example, consonants interacting in complex (branching) onsets, or in coda–onset contacts, and vowels interacting with the following codas in closed syllable effects. Thus, with CV as the basic representational grid the main theoretical questions are what makes the different configurations in (3) grammatical, and what makes them behave the way they do in the new mode of speech sound organization? How are the empirical effects captured earlier by arboreal structure translated into the Strict CV model?

Given that in Strict CV all syllables are formally open, and all clusters in fact involve heterosyllabic consonants, the question is what allows some clusters, typically RT, to close the preceding syllable and preclude length of the preceding vowel, as in e.g. *reception*, and what makes vowel length possible in front of some TR clusters (branching onsets), e.g. in *cobra*? Also, what is the source of variation in the latter context? These questions are not trivial. In traditional models, it was the syllabic grouping of phonological strings that was made responsible for such effects. It is clear that the CV representation on its own is unable to address these questions and some mechanisms must be recognized, which are responsible for the syllabic effects known by such names as closed syllable shortness, open syllable lengthening, metrical lengthening, extrasyllabicity, coda effects, positional strength,[5] vowel-zero alternations, phonotactics, etc. This is precisely what CVCV is all about. It shifts the labour from arboreal to lateral structure which is due to a network of lateral relations established after morphemes are concatenated (cf. section 1.2).

Various incarnations of the Strict CV model, unlike Standard GP, are also strict on the number and type of computational mechanisms that may be part of phonology, recognizing only two familiar mechanisms of government and licensing as major organizing principles. The two lateral forces now have a precise definition in terms of what they do to their targets. But the ways in which these lateral relations are employed may differ across CV frameworks, and make markedly different predictions.[6] Below we look at a fairly well-established development of Strict CV called Lateral Theory of Phonology (LTP) in its most recent version.[7]

10.2.3 LTP with Coda Mirror 2 – main assumptions and predictions

Exactly how the C and V positions interact with each other to yield the observed syllabic effects is a matter of some debate, which is why we use one model to demonstrate what predictions such a restrictive phonological system can make. One way in which all versions of CVCV differ from Standard GP is in complete flattening of the phonological representation and replacement of traditional syllabic (arboreal) constituents with lateral relations between positions. Another difference lies in the elimination of the confusion as to what the two lateral forces can do. Government is assumed to be a relation that diminishes melodic material under the affected position. Thus, it may be viewed as a negative or destructive force. Licensing, on the other hand, supports melodic structure. It is a positive force. This distinction was not so obvious in Standard GP in which a governed position was said to be also licensed. Additionally, all lateral relations in LTP are right-headed. This follows from Standard GP directionality, which is regressive among constituents (ICG, as discussed in section 10.1.2).

10.2.3.1 Vowels as a source of government and licensing

Lateral relations of government and licensing are by default discharged by melodically filled nuclei (vowels), which are sometimes called lateral actors. It is also assumed that both lateral forces are universally present if a given nucleus is an actor. Also under special conditions empty nuclei may be lateral actors. This concerns mostly the word-final empty nucleus (FEN), which may be given the status of a lateral actor by means of a systemic parameter. The internal empty nucleus (IEN) may also be a lateral actor, but only if it is sandwiched within a cluster of rising sonority (TR), which involves an infrasegmental relation (see (9a) below). We begin by looking at lateral actorship of vowels and will extend the discussion to empty nuclei as actors in due course. The configurations in (4) show the targets of government.

(4) Targets of government

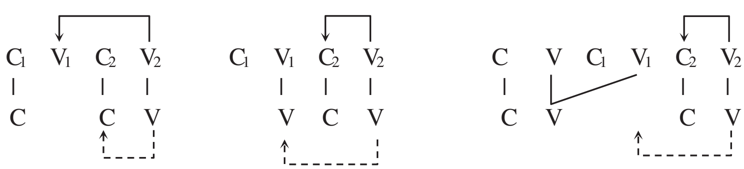
a. empty nuclei b. intervocalic onsets

Government discharged from a vowel always falls on the preceding V or C. The choice depends on whether the preceding nucleus is empty, as V_1 in (4a), and requires government as per ECP, or not. If no empty nucleus precedes, government from V_2 is exhausted on the adjacent ONSET C_2 (4b).

The configurations in (5), on the other hand, show targets of licensing. By convention, licensing relations are shown by dotted arrows underneath the structures.

(5) Targets of licensing

a. onset b. nucleus c. nucleus



There are two main principles governing the distribution of the two lateral forces. Firstly, no position is both licensed and governed at the same time. And secondly, government takes precedence. It goes to the preceding empty nucleus (5a), or to the preceding ONSET if the preceding nucleus is not empty (5b,c). Licensing, in turn, affects the position that escaped government. The licensing relation in (5c), in which we see a long vowel, is crucial in allowing length. Thus, the presence or absence of this licensing relation is directly responsible for the open vs. closed syllable effects mentioned earlier (see section 10.2.3.3 below for more details). Note, that the domain of application of government and licensing, namely, the preceding VC, involves the immediately adjacent preceding ONSET, and the immediately preceding nucleus, where the latter relation takes place at a nuclear projection. Thus, the lateral relations are local. Being local means that no V intervenes between the two nuclei related by a lateral relation.

10.2.3.2 Strong and weak positions in Strict CV, Coda Mirror

Formally defined positional strength of consonants determines their propensity to such phenomena as lenition or fortition. Given that relative strength may vary depending on position in the word, we expect from a theory to designate precise cut-off points for the scope of particular processes. To be more precise, a theory should be able to predict which positions of varying strength may be affected by a given process and what implicational relationship between positions is at play.

The configurations of the destructive (government) and supporting (licensing) forces make clear predictions as to the positional strength of segments. Below we replicate a summary of cross-linguistically observed strong and weak positions from Scheer & Ségéral (2008b: 486). Recall that traditional syllable-based generalizations concerning prosodically strong and weak positions, which base the distinction on the segment's location in the 'coda' for weak and in the 'Onset' for strong, fails to capture the fact that an intervocalic Onset is nonetheless weak (6e). In fact, even the strong positions cannot be described by regular syllabic inventory. Not only are intervocalic onsets weak but also there is clear variation as to the strength of the word-initial consonants which may be strong or not. Additionally, there is a much ignored fact that post-coda consonants are strong. We intend to show how the disjunction {#,C}__ can be captured in Strict CV, allowing also for variation in the word-initial context.

(6) Strong and weak positions

	position	usual name		
a.	***#_V***	word-initial	strong positions	
b.	VC._V	post-coda		
c.	V_.CV	internal coda	coda	weak positions
d.	***V_#***	final coda		
e.	V_V	intervocalic		

What needs to be noted about (6) is that two contexts exhibit cross-linguistic variation. The word-initial consonant may or may not exhibit strength, while the word-final consonant may or may not exhibit typical coda effects. The two contexts 'Onset' (6a) and 'coda' (6d), which happen to occur at word edges, will be discussed in detail shortly.

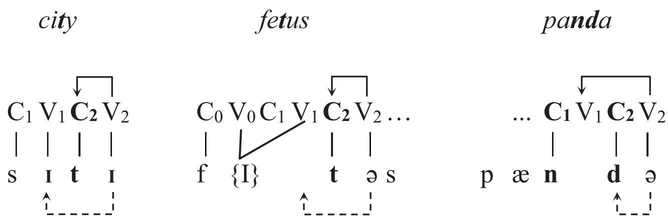
Let us now take a detailed look at how some of the empirically observed patterns with respect to positional strength from (6) correspond to the various configurations of government and licensing defined above in (4) and (5). We temporarily omit the left edge of words (6a), as well as clusters with rising sonority (TR). However, we introduce the FEN which is parametrically set to be a lateral actor in, for example, English.

The positional prosodic strength of consonantal segments is directly deducible from the arrangement of the lateral forces. In the most recent version of LTP, a position can be governed, licensed or unaffected by any of the forces. Beginning with (7a), we see that V_2 is a melodically filled nucleus and therefore a lateral actor. Since V_1 does not call for government because it is not empty, this lateral force is exhausted on the Onset C_2. Licensing, in turn, cannot land on a governed position, therefore it goes to V_1. The governed C_2 is in a prosodically weak position, in which we expect lenition. Additionally, the arrangement in (7a) defines open syllables in LTP: a syllable is open if its nucleus is licensed. This will have consequences for vowel length, which typically requires licensing, as in the first vowel in *fetus* [fiːtəs].

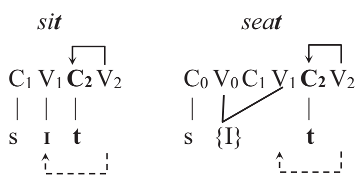
(7) Strong and weak positions in English: three degrees of positional strength

a. *C2 – intervocalic: weak*

b. *C1 – internal coda: neutral*
C2 – Coda Mirror: strong



c. *C2 – word-final = intervocalic: weak*



In (7b) we are dealing with a 'coda–onset' contact of falling sonority (RT cluster). The two consonants are separated by an IEN, which in a sense determines their respective positional strength. Beginning with C_1, we observe that it is followed by a governed empty nucleus which, for this reason, is not a lateral actor itself (hence, no arrows stem from V_1). Therefore, C_1 is neither licensed nor governed. Note that this consonant corresponds to what is traditionally called an internal coda. Formally, it finds itself in a better position than the intervocalic consonants in (7a). Whether this prediction is correct remains to be seen, and some discussion will be offered shortly. It should be noted, however, that we can now provide a formal definition of what a 'coda' consonant is. It is an Onset followed by a governed empty nucleus, i.e. a nucleus which is not a lateral actor: it cannot govern or license. We will see the consequences of that shortly.

Turning now to the second consonant in the RT cluster in (7b), it must be said that the post-coda consonant is prosodically strong (cf. 6b). This correlates with, for example, its diminished propensity to lenition. The question, however, is how this fact can be captured formally. Recall that pre-vocalic context is insufficient in defining strength of onsets, because intervocalic onsets are pre-vocalic and they are weak (6e). Note also that C_2 is not governed by its nucleus because government is needed to sanction the IEN in V_1. Therefore C_2 is licensed. Ségéral & Scheer (2001) and Scheer & Ségéral (2008b) call this position Coda Mirror, which has a double meaning. Firstly, unlike codas, it is strong. And secondly, unlike codas, which precede a governed empty nucleus, the consonants in Coda Mirror follow one. Thus the pre-theoretical mirror: _{#,C} vs. {C,#}_, corresponds to a mirror effect, i.e. that the opposition in behaviour (strong–weak) can hardly be accidental given the mirror in the structural description. The presence of the governed empty nucleus in RT clusters is crucial. It deflects government from C_2, which in turn is licensed and strong. Below, we will see that the Coda Mirror position may also occur word-initially and may be responsible for the strength of word-initial onsets in some languages.

Let us now turn to the last two representations in (7c). The new aspect here concerns the FEN in V_2, which is parametrically allowed to be a lateral actor in English. There are a number of diagnostic criteria allowing us to establish this parameter setting. For example, the existence of word-final clusters in words like *bend*, *act* tell us that the FEN can govern the intervening empty nucleus in the final cluster. Similarly, the presence of long vowels followed by one word-final consonant is only possible if the FEN is a lateral actor, e.g. *beat*, *wise*, because the long vowel requires licensing. Thus, given that the FENs in (7c) are lateral actors, the configurations are formally the same as in (7a) in that C_2 is governed. The relevant consonant C_2, therefore, behaves as if it were intervocalic with respect to its positional strength. The difference between (7a) and (7c) is substantive: the final nucleus is or is not melodically filled, which influences the types of lenition that may take place in the two respective contexts.

10.2.3.3 Long vowels, closed syllable shortness and extrasyllabicity

Since vowel length is licensed by the following nucleus, we must look at such effects as closed syllable shortness and extrasyllabicity from a different perspective. Consonants do not close syllables. It is the absence of internuclear licensing that does it.

Above, we discussed some examples of how government and licensing can determine the positional strength of consonants. Here, we look at how the same lateral forces determine what effects we should expect on the targeted preceding V. Recall that the preceding V is governed if it is lexically empty (7b), and licensed if it has melody (7a,c). Of the main two effects of internuclear relations, that is, vowel-zero alternations and vowel length effects, we only concentrate on the latter now.[8]

As we saw in (7a,c), vowel length requires licensing from the following nucleus. In fact, the so-called open syllable is defined as one in which the nucleus is licensed. It follows that no length will be possible if that licensing is absent. In general, this takes place in two situations.[9] Firstly, such licensing is missing if the following nucleus is empty and governed, as in *panda* (7b). Since the IEN V_1 is not a lateral actor, the preceding nucleus cannot be licensed, and cannot be long. This situation is what we refer to as closed syllable shortness, which may have different concrete outcomes. It refers to a static situation, but also to closed syllable shortening and impossibility of tonic lengthening. The second situation in which licensing may be missing is at the right edge of words, when a consonant is followed by a FEN which is systemically set not to be a lateral actor. Since such FEN neither governs nor licenses – precisely as the internal governed empty nuclei – the word-final consonant will behave like a coda. In other words, the consonant will not be extrasyllabic like in English *feet*, and we expect closed syllable shortness effects mentioned above. Note that in this model, the actual explanation of this effect does not really refer to the consonant but to the type of relation that the FEN contracts with the preceding nucleus. It is not that the final consonant acts (non-extrasyllabic) or does not act (extrasyllabic) like a coda. Extrasyllabicity in LTP terms means that the FEN is an actor, and the preceding C is thus in an intervocalic position, that is, governed. This definition is important when we want to understand the variation related to word-final consonants, which may behave as codas or not (Piggott 1999; Scheer 2012a), as opposed to word-medial coda consonants, which invariably close the preceding syllable. This is because internal codas are followed by an IEN, rather than by a FEN, and only FENs are a locus of variation: morpheme edges are subject to parametric influence. Thus, extrasyllabicity is no longer an extraneous mechanism attempting to capture the unexpected invisibility of final consonants. Rather, it is a result of a parametric choice to grant FEN actorship or not, neatly corresponding with the empirical fact that in some languages the final coda does behave as if it were extrasyllabic, while in others it does not.

Let us illustrate word-final closed syllable shortness with two examples, which in fact show interesting additional variation that Strict CV can cover. Kaye (1990), in his article defending the proposal that all words ending in a consonant in fact end in an empty nucleus, used two workbook examples of languages exhibiting closed syllable shortness (Turkish and Yawelmani) both word-medially and word-finally. We mention these two cases because they both differ from English, which exhibits word-final extrasyllabicity, but they also differ from each other in an interesting way. We take the same examples as Kaye (1990: 302) for simplicity. Kaye identified the precise context for shortness in these forms as occurring before an empty nucleus in the following syllable.

(8) Yawelmani

Aorist Pass.	Aorist	Fut. Pass.	
sa:p-it	sap-hin	sap-nit	‘burn’
pana-t	pana:-hin	pana:-nit	‘arrive’

Turkish

NOM.	POSS.	ABL.	NOM. PL.	
merak	mera:k-ɨ	merak-tan	meraklar	‘law’
sevap	seva:b-ɨ	sevap-tan	sevaplar	‘good deed’

Beginning with Yawelmani, and applying the lateral forces of LTP, [panat] < /pana:+tø/ is a case of shortening a lexical long vowel in front of FEN, while [saphin] < /sa:pøhinø/ shows shortening before a medial governed nucleus (IEN). We may conclude that both IEN and FEN are not lateral actors in Yawelmani. That this is correct can be demonstrated by the fact that this language does not allow word-final clusters or clusters of three consonants. It is interesting that the rules for vowel shortening and vowel epenthesis are identical in terms of the following triggering context (cf. shortening: VV → V/_ {C#, CC} and epenthesis: ø→ V/C _ {C#, CC}). Note that CC# translates into Strict CV /CøCø#/, that is, a sequence of IEN and FEN. The two separate rules follow from the same cause: the FEN and IEN are not lateral actors in Yawelmani. Therefore, vowel length is not licensed, resulting in ‘closed syllable shortness’, and IEN cannot be governed by the FEN, so it has to be realized phonetically (epenthesis).

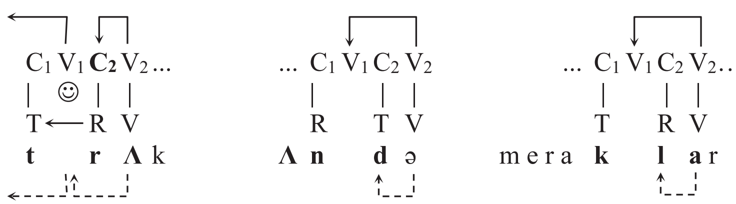
Turkish appears to work in an identical way, except that it has word-final clusters of falling sonority (RTs), e.g. [sarp] ‘steep, nom.’. How is this possible? The presence of these clusters suggests that the FEN in fact is an actor and it can govern the preceding IEN. If so, then the question is what causes the ‘closed syllable shortness’ in [merak]? The paradox is only apparent. It is an established fact in GP that empty nuclei are weaker licensers than nuclei with melody (Charette 1990; Harris 1994; Cyran 2010). Thus, all we need to say is that the FEN in Turkish is not a strong enough licenser to support vowel length in the preceding nucleus. In other words, formally speaking, Turkish is like English in that FEN is a lateral actor. Only the licensing strength of that type of nucleus does not allow for licensing vowel length. This situation is not special in any way. A similar situation is observed in dialectal English. In the so-called Aitken’s law in Standard Scottish English FEN is unable to license vowel length in the preceding syllable depending on the nature of the intervening consonant (e.g. Zdziebko 2012). Thus, both in Turkish and in Yawelmani the FEN cannot license long vowels, but for different reasons. In the former, the FEN is an actor but a weak licenser, while in the latter, it is not a lateral actor at all.

10.2.3.4 Complex onsets (TR)

Rising sonority clusters (TRs) exhibit varying behaviour with respect to the preceding vowel. They either act like single consonants leaving the preceding syllable open and allowing for vowel length, e.g. in English, or they close the syllable like RT clusters, e.g.

in Turkish and generally Semitic languages.[10] The former type of TR corresponds to the traditional branching Onset and has a special structure in LTP, involving Infrasegmental Government (IG) (Scheer 1996, 2004). It is a relation between a sonorant and the preceding obstruent (9a). This relation allows the intervening IEN to remain empty and be a lateral actor, which we mark with the happy face symbol ☺ (see below). The spurious TR clusters, on the other hand, result from government of the intervening IEN (9c). In this sense, they are formally identical to RT clusters (9b), and they behave like RT. This is visible in, for example, Turkish [mera:kɨ ~ meraklar] 'law, nom.pl.', where the vowel is shortened before TR. True branching onsets allow for vowel length as in Italian [pi:gro] 'lazy', or Icelandic [nɛ:pʰja] 'cold weather'. The representation of the branching Onset below incorporates the proposal of Brun-Trigaud & Scheer (2010) into Coda Mirror 2, whereby V_1 is a lateral actor.

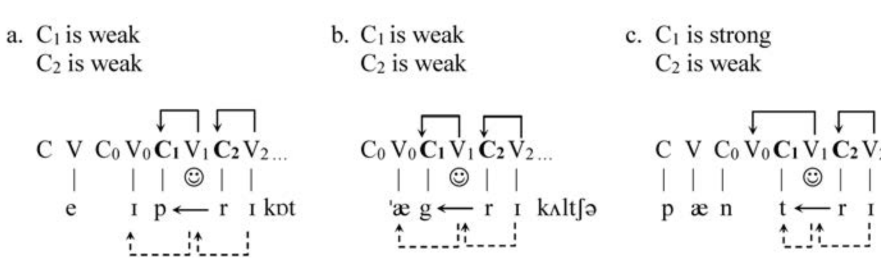
(9) a. branching onset b. coda -onset c. spurious TR



One important difference between the fate of V_1 in (9a) and (9b,c) is that the former is not governed. The ECP is satisfied in (9a) by the fact that the surrounding consonants have contracted an infrasegmental governing relation.[11] Since that relation sanctions the IEN's emptiness, V_1 does not call for government from the following nucleus, and that lateral force is expended on C_2. Thus, V_1 in (9a) is in fact licensed by V_2 like any ordinary vowel in open syllables. This in turn means that it is itself a lateral actor. The arrows of government and licensing are deliberately not targeted at any position, but it is obvious that their arrangement will depend on what precedes, just as we saw with single consonants in (7). Details will come shortly. The important thing to note, however, is that in languages which do not allow for Infrasegmental Government, surface TR clusters will have the structure as in (9c) and will behave in the same way as RTs in (9b), namely, the V_1 will be governed and will not be able to act as a lateral actor itself. This will have consequences, for example, for the distribution of long vowels.

Below, we look in more detail at various configurations involving TRs in English with special focus on the positional strength of the two consonants. We look at the same contexts, that is, following a long vowel, a short vowel and an empty nucleus.

(10) Branching onset in different post-nuclear contexts in English

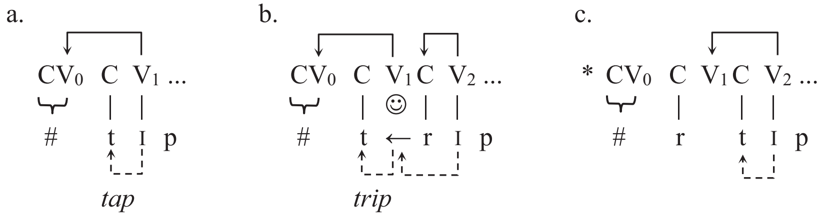


In all the structures in (10) the relevant empty nucleus V_1 is made silent by Infrasegmental Government and does not require government. This force is directed to C_2, while V_1 is licensed and, like any ungoverned nucleus, it is a full-fledged lateral actor that can govern and license. Note that the sonorant in C_2 is always in a prosodically weak context (intervocalic), that is, it is governed but unlicensed. This may be the reason why this position in branching onsets is so restricted melodically. The situation of C_1 changes depending on the nature of the preceding nucleus. In (10a,b) C_1 is weak. It is governed because the preceding nucleus does not require government, and in return that nucleus receives licensing. This is what makes TR clusters behave like single consonants in some languages in that this cluster does not cause closed syllable shortness. Note that the fate of C_1 is different only in (10c). In fact it is in Coda Mirror position: it is licensed but ungoverned. Thus we should expect weakening phenomena in TRs in (10a,b) as opposed to (10c). This prediction seems to be borne out in the way Latin TR clusters developed into Modern French, as shown in Brun-Trigaud & Scheer (2010). The obstruents of the intervocalic TRs were lenited, while post-coda ones were not, for example, Lat. *ca**pr**a* > Fr. *chè**vr**e*, but Lat. *com**pr**end(e)re* > Fr. *com**pr**endre*. There is a general shortage of empirical studies and information on the specific behaviour of TRs in lenition. It is only in CVCV that attempts have been made to come by stable cross-linguistic data. What appears from the French case, but also from the other phenomena quoted in Brun-Trigaud & Scheer – Celtic and Gorgia Toscana, but also the dialectal evidence from French – points to the following pattern: T in TRs is strong after consonants and weak after vowels, exactly as singletons – as if there were no R.

10.2.3.5 At the left edge of words

We saw that the relationship between government and licensing in different configurations allows us to make certain generalizations concerning the phonotactics at the right edge. The situation is no different at the left edge.

Lowenstamm (1999) proposed that major syntactic categories begin with an empty CV. In the LTP version of Strict CV the empty CV site is parametrically present or absent and the decision may be limited to particular syntactic structures, for example, phrase level (Scheer 2012b). Since the initial CV site has to be phonologically interpreted – the empty nucleus calls for government – the presence or absence of the initial CV divides linguistic systems into two empirical situations with respect to word-initial phonotactics. Systems like Polish, which do not have the initial CV, enjoy relative freedom as to the sonority profiles of initial clusters. Words can begin with TR, RT, RR and TT, e.g. Polish *trawa* 'grass', *rtęć* 'mercury', *lnu* 'flax, gen.', *kto* 'who'. It is because initial clusters that contain an empty nucleus requiring government will be grammatical. This is not true of systems which possess the initial CV, as can be seen in (11) below. English, which has an initial CV site, can only allow for clusters of rising sonority (TR), that is, the true branching onsets with Infrasegmental Government (11b). The initial consonant in such systems is strong, as per Coda Mirror (11a,b), while initial RT is banned (11c).

(11)


The empty nucleus V_0 can only be governed in (11a) and (11b). In (11c) V_1 is not a lateral actor because it is governed by V_2. The structure with an ungoverned V_0 is ungrammatical. It is crucial to note that the absence of initial RT clusters in languages like English is caused by exactly the same principles that rule out vowel length in closed syllables. Any RT contains an empty nucleus which requires government from the following vowel. This excludes the possibility of having another empty nucleus in front of RTs. In Polish, on the other hand, clusters like (11c) are possible because there is no initial empty CV. There are other effects and predictions associated with the presence of the initial CV. Scheer (2004) connects the stability of the first vowel with their role as governors of the initial CV, as well. Thus we are dealing with a convergence of seemingly unrelated phenomena at the left edge, which stem from one aspect. As for predictions, the model makes it clear that #TR-only languages, which possess the initial CV, must also show strength of single consonants, e.g. English, and vice versa. #Cs are weak in anything-goes languages, e.g. Greek.

To sum up, the parametric presence of the initial CV site allows us to make a distinction between languages with relative phonotactic freedom at the left edge and those with a very strict shape at the left edge. Additionally, we identify the reason why word-initial position is strong in some languages. This fact goes hand in hand with the above mentioned phonotactic patterns, and the initial strength is defined in exactly the same way as post-coda strength. It is due to Coda Mirror. The strong consonant is licensed. Thus, what is captured in this model is not only the parametric nature of initial strength (6a) but also the disjunction {#,C}__, which is neatly reduced here to one causality, something that regular syllable structure is unable to do. It is also important to note that the two contexts form an exact mirror of the coda (__{C,#}) and have the exact mirror effect: strength as opposed to weakness.

In conclusion, it must be stressed that, unlike in many syllable-based models, in Strict CV lateral forces do two jobs at the same time: they define structure (who is a coda, who is an Onset) as well as designate strong and weak positions. In regular models, the theory of syllabification remains unmodified, while whether codas could be strong or initial consonants weak is a matter of a separate theory of lenition. In Strict CV, the fate of both aspects, that is, structure and causality, is connected. As a result, Strict CV makes predictions that others do not make.

10.2.4 Some challenges and current issues

10.2.4.1 Positional strength and contradictory phenomena

Recall that the following types of nuclei can be actors: (a) filled nuclei by default, (b) FEN by parameter and (c) IEN inside TRs, of which the last one is the most special and rare situation. The typical non-actors are (a) FEN by parameter and (b) governed IEN (in RTs and bogus TRs). Thus, the generalization is that all governed nuclei and only those are non-actors, while all ungoverned nuclei and only those are actors. With respect to positional strength, there are only three types of configuration in which a consonantal segment can find itself (12). Note that we are now in a position to explain the variation at word edges mentioned in (6a) and (6d). On the left edge, the variation comes from the parametric distribution of the initial CV, while the right edge variation is due to the parametric decision whether FEN is a lateral actor. The existence of variation at edges, as opposed to non-variation morpheme-internally is an important empirical generalization, which goes unnoticed in the non-GP literature: the word-internal piece of the mirror conjunctions __{#,C} and {#,C}__ is invariable, while the one touching the edge is variable.[12]

(12)

a. governed (weakest) –
   i. intervocalic — ***city*** [sɪʔɪ]
   ii. final single C when FEN=actor — ***sit*** [sɪʔ]

b. unlicensed and ungoverned (neutral/weak) –
   i. internal coda — *wi**t**ness* [wɪʔnɪs]
   ii. final coda when FEN≠actor — [panat] in Yawelmani[13]

c. licensed (strongest) –
   i. post-coda — *pan**d**a* [pændə]
   ii. word-initial — ***t**ip* [tʰɪp]

A comment is in order concerning (12a) and (12b). Often these two formal positions pattern together in phonological processes, even though they are two disparate contexts in LTP. One example where the two contexts act uniformly is in English t-glottalization. The phenomenon may be found in intervocalic (*better* [beʔə]), final (*bit* [bɪʔ]) and internal (*chutney* [tʃʌʔnɪ]) coda contexts. What unites these contexts is only the fact that they are not strong, that is, licensed. There are, however, processes which distinguish the intervocalic onsets from the final and internal codas. This is a challenge for this model because the effect is found in context (12b) and only in half of (12a). Namely, it does not occur intervocalically. Not only that: the two contexts in which it does occur are not equal. This concerns, for example, the allophonic distribution of clear and dark /l/ in English. Recall that we find dark [ɫ] word-finally and in front of a consonant, but not intervocalically. Clearly we are dealing with some other conditioning here; something to do with the presence or absence of melody in the following nucleus.[14] Note that word-final dark [ɫ] may become clear [l] in front of vowel-initial words, e.g. *fail* vs. *fail it*, or *oil* vs. *oil on troubled water* (Gussmann 2002: 12). It is therefore quite possible that the lateral allophony has nothing to do with government and licensing relations established inside words, unless we postulate that such relations are also established across words in English. What this means is that the lateral allophony could be non-phonological in nature.

Another process which seems to cut across the distinctions in (12) in the same, albeit reversed, way is t-tapping in English. It occurs intervocalically, excluding the formally identical context before FEN. So again, here the distinguishing factor is presence or absence of melody in the following nucleus and not the formal configuration that follows from the model. Admittedly, these conclusions hang on the assumption that the FEN in English is a lateral actor. There are good reasons to believe that the assumption is correct. Firstly, final consonants are extrasyllabic and allow for preceding long vowels, which, when translated into LTP, means that long vowels are licensed by the FEN. Secondly, English exhibits word-final clusters, e.g. *act*, *belt*, *lisp*, etc., which suggests that the FEN governs the empty nucleus inside these clusters. Thus, the status of the FEN in English as actor can be established independently of the lenition phenomena, and on the basis of stronger diagnostics than lenition. This in turn forces us to reconsider the status of some of the lenitions, or to rethink the relation between nuclei and their onsets.

Examples of the sort mentioned above can be multiplied, and the solution seems to be at hand. Some licensing relation needs to be recognized between onsets and their nuclei regardless of the arrangement of lateral forces in current LTP. For example, it may be the case that we need to go back to the Standard GP distinction between prosodic (p-licensing) and autosegmental (a-)licensing. And only p-licensing is the lateral force interacting with government, while a-licensing is always present within each CV pair and takes into account the melodic shape of the nucleus. This possibility must be further studied, as some formal analysis needs to be proposed for cases where not only the arrangement of the lateral forces but also the nature of the nucleus is included. Note that at this point the intervocalic and final single consonant is governed in languages where the FEN is an actor, and it is in no licensing relation with it. Thus, the relation between this consonant and the following nucleus with or without melody is at this point not expressible in phonological/computational terms. This still leaves the possibility that some phenomena, for example, clear [l] in pre-vocalic sandhi contexts, may be interpretational rather than phonological.

10.2.4.2 Licensing with no government?

One possibility which is worth considering is that we are dealing with one lateral force too many, an option that seems to be valid at least for languages with no clusters. Kula & Marten (2009) question the assumption that lateral actors must universally both govern and license. They attempt to demonstrate on the basis of Bantu languages that clusterless languages also do not use the initial CV. Thus, the two prime reasons for having government (no empty nucleus in initial CV site and none inside clusters to govern) are missing in such languages. This leaves the question as to what accounts for the positional strength of initial consonants and positional weakness of intervocalic consonants in such systems. Kula and Marten propose that since there is no government in such systems, the first vowel in the word licenses its Onset because it is not governed, and so this Onset is in exactly the same situation as the one in languages with government sanctioning the initial CV. A Coda Mirror effect for the initial context can be therefore achieved without government.

The relative intervocalic weakness of consonants in clusterless languages, on the other hand, is assumed to follow from Licensing Inheritance known from Standard GP (Harris 1997). The suggestion is that every nucleus receives its licensing potential from the following one. Intervocalic consonants are licensed by vowels that need to also license the preceding nucleus, which is not the case word-initially because there is no preceding nucleus to license. Therefore, initially, the consonant gets all the licensing while intervocalic onsets share licensing received from the following nucleus with the nucleus that precedes them. That licensing seems to be 'divided' for both positions in the preceding VC has also been noted in English (Zdziebko 2012) and Polish (Cyran 2014a).

There are a few problems connected with this proposal. It is not clear what would happen if a clusterless language is found to use an initial CV site, whose presence would be indicated, for example, by morphological considerations, or phenomena unrelated to lenition, e.g. gemination of the consonant. Government would also need to be 'activated' in cases of diachronic evolution where vowels start to alternate with zero and fall out, creating clusters. The licensing-only proposal also makes a prediction that the initial consonant in clusterless languages is always strong, while LTP predicts that it may be strong or weak, depending on the distribution of initial CV. The model with no government also does not extend easily to languages possessing clusters in which the initial as well as post-coda strength indeed depend on the presence of the government with respect to the preceding empty nucleus.

Another alternative to LTP, which in a sense provides an answer to how government is activated in systems with clusters, is one in which licensing is a lateral force that is expended only by nuclei, in fact, all types of nuclei, while government is limited to interconsonantal relations. An example of this option is Cyran's (2010) Complexity Scales and Licensing model (CSL), which in many respects translates Standard GP governing relations from obstruent to sonorant (T→R, R←T) into a CVCV framework. The governing relations are subject to Government Licensing (Charette 1990) in that each such relation must be licensed by the following nucleus. The interconsonantal government 'locks' intervening empty nuclei and makes them invisible phonologically (cf. Szigetvári 1999). Cyran also eliminates internuclear government, reducing vowel-zero alternations to the operation of the NoLapse constraint (*ø-ø) disallowing sequences of two visible empty nuclei (Rowicka 1999).

One clear advantage of CSL is that it is able to formally express the implicational relationship between RT and TR clusters in languages (see section 10.2.4.4 below), but the elimination of internuclear government and the invisibility of empty nuclei inside RT clusters makes this model incompatible with the findings of Coda Mirror. Positional strength must somehow be redefined, possibly along the lines of Harris (1997). One aspect, however, which it emphasizes, and which seems to be lacking in current LTP, is that every active nucleus, whether filled or empty, must be in some licensing relation with its Onset, albeit only in terms of autosegmental rather than prosodic licensing.

10.2.4.3 Direct and indirect effects of government and licensing

The two lateral relations in LTP are responsible for a whole range of phonological phenomena including static phonotactic patterns as well as observable phonological processes. One of the direct effects of government and licensing is vowel-zero alternation, which may be understood as epenthesis or syncope depending on the system. The force that is responsible for these phenomena is internuclear government. Another directly observed phenomenon is vowel length alternation, which can take the form of closed syllable shortening of lexically long vowels (Turkish, Yawelmani, English) or tonic lengthening (Italian, Icelandic). This set of phenomena is directly due to licensing, but also indirectly to government, in the sense that governed empty nuclei (IEN) do not provide licensing that is required by long vowels. Another set of phenomena which are directly observable and which are due to the arrangement of government and licensing are all sorts of weakening processes. These are, as shown above, typically related to the distinction between filled nuclei and empty nuclei and their relation to the preceding Onset (e.g. t-glottaling, t-tapping, l-allophony, to name but a few familiar phenomena from English, but also voicing alternations in e.g. Polish).

The indirect effect of the arrangement of the lateral forces mainly lies in the fact that the forces define relative positional strength of consonantal positions and so determine the probability of, for example, lenition processes in particular positions as well as make predictions as to the implicational relationships between these effects. Exactly what is lenited and in which set of non-strong positions as opposed to non-weak positions is a systemic decision. Positional strength, however, is not absolute, but relative. It is predicted that lenition also occurs in strong positions. What is systemic is the decision concerning when something lenites, what lenites, how it lenites and in which positions – the choice of the position, though, is universally constrained by an implicational hierarchy defined by positional strength. If a process lenites an item in a stronger position, the same item will also be lenited at least as much in all weaker positions. Hence voiceless stops affricate in the high German consonant shift in strong position, and they lenite more (to fricatives) in intervocalic and coda position. The same goes for fortition: if fortition occurs in a position X, it will also occur in all stronger positions. The actual effects are additionally constrained by the nature of the melodic representation used in LTP. Thus, for example, synchronically speaking, positional fortition which does not involve spreading of elements from the surrounding context can only be interpretational, that is, a case of a shift in phonetic interpretation of the same representation, and not computational, that is, phonological. This is due to how Element

Theory is constructed, which precludes non-local sources of melodic categories. Changes in the interpretation of objects in such contexts may be followed by lexicalization/phonologization. Only in this diachronic sense may elements be added from 'nowhere' to strengthened objects (see Chapter 11, or Cyran 2014b).

10.2.4.4 TR vs. RT

One hitherto unexplored linguistic pattern within LTP is the phonotactic relationship between rising sonority (TR) and falling sonority (RT) clusters across languages. There is a well-established implicational relationship between the two types of clusters which goes back to Kaye & Lowenstamm (1981), whereby the presence of TRs seems to be more marked and implies the existence of RTs in a given system, but not the other way round. This relationship is particularly visible in word-final position. Languages like English or Turkish do not allow for TRs finally, while accepting RTs. Polish and Icelandic, on the other hand, possess both types of clusters. Above, we saw that in Turkish word-internal RTs and some TRs involve a governed empty nucleus, which should make these structures formally equal, in the sense that it does not matter what the melodic profile of the consonants surrounding the governed IEN is. However, if LTP may be modified a little to accommodate the fact that all onsets are in some licensing relation with their nuclei, albeit only autosegmental, and that the various types of nuclei exhibit varying onset-licensing strength, then even the formally identical RøT and TøR become different. All that needs to be assumed, like word-finally, is that when first introduced into the system, empty nuclei are weak onset-licensers. If so, given that obstruents are more complex representations than sonorants in terms of the number of elements, it is understandable that it will be easier to license a sonorant in RøT than an obstruent in TøR. It should be emphasized, however, that we are not talking about the prosodic licensing which is targeted at different positions depending on its alignment with respect to government, as standard LTP would have it. For the purpose of this discussion we call this 'new' and additional licensing 'autosegmental', for want of a better term, and it is present in each CV. It is a matter of further study what this relation is, and whether LTP can in fact do without it.

Recall that true TRs additionally involve a relation of Infrasegmental Government T←R, and as a result behave as if they were a single consonant in that they keep the preceding syllable open. This additional mechanism may be responsible for the fact that true TRs are even more marked than RTs and bogus TRs. Hence their restricted distribution word-finally, when followed by a FEN – clearly a weaker licenser than a melodically filled nucleus. In other words, the universal phonotactic patterns involving RT vs. TR are to a great extent predicted in LTP, but some modification of the system is required, namely, we need to admit that there is no such thing as a non-actor nucleus. Every nucleus is an actor of sorts, but its lexical representation and its place in the network of lateral relations determines its varying licensing potential.[15]

10.2.4.5 At the right edge

As we have just seen, LTP is able to capture quite a range of phenomena at the right edge of words with just the two lateral forces: licensing and government, which organize phonological representation in all positions. The parameter designating the FEN as a lateral actor makes a very interesting prediction concerning the right edge of words. If the FEN is not an actor, this immediately means a few things. Firstly, the preceding nucleus will not be licensed or governed. This means that the preceding nucleus is not in an open syllable situation and vowel length is not possible before word-final single consonants (Yawelmani). Secondly, the final C finds itself in a relatively neutral position (neither strong, nor weak), that is a typical coda. And thirdly, the absence of the lateral force of government disallows word-final 'complex codas', that is, RT clusters, which must contain a governed IEN as in English *bend* /benødø/ and *act* /ækøtø/, or Turkish [sarp] 'steep, nom.'. Thus, a system with a non-actor FEN may end its words in one consonant which closes the syllable just as the internal codas do. This system may have clusters and long vowels, but only before filled nuclei in the following syllable.

The positive setting of the parameter on FEN actorship makes radically different predictions. Firstly, word-final single consonants are extrasyllabic and allow for preceding long vowels. The word-final single consonant is extra weak because it is governed, as any intervocalic consonant. Additionally, this system allows for word-final RT clusters like English *act* and *bend*, in which the last consonant is in a strong position because it is licensed. And this is where some problems appear to arise concerning static phonotactics of the right edge. Of course it should be borne in mind that LTP does not aim to directly express static patterns. Rather, it makes predictions for where and how phonological processes such as lenition and fortition are likely to occur. However, it is also true that some static patterns have resulted from historical processes whose application we do predict. Thus, returning to static patterns at the right edge of words, we predict that systems with FEN actorship will have robust RT clusters – with the T in strong Coda Mirror position (licensed) – and very weak single final consonants (governed). Additionally, we predict that word-final single consonants (simplex codas) in a language like English are prosodically weaker than the internal codas. This implication follows from the fact that in a system in which the FEN is an actor, the final consonant is governed, while the internal coda is neither licensed nor governed. It remains to be seen if there is an empirical echo of these predictions. In some cases, however, it appears that the opposite seems to be true. Namely, that some lenitions affect internal codas, but not final C#. One case in point might be the retention of [l] in word-final context in the shift from Latin to French, e.g., *sal* > *sel* 'salt', and l-vocalization in internal codas, e.g., *alba* > *aube* 'dawn' (Scheer 2012a). According to LTP, final C# can be equal in strength to internal codas, but never stronger.

Again, it seems that these apparent problems of the theory could be overcome if we assume that there are no 'naked' consonants and that each CV is a licensing relation regardless of the type of nucleus. Under this assumption we could say that empty nuclei can only ever be used in a system if they are allowed to license some melody in the preceding Onset. Otherwise they are not used.

All in all, the recurring theme in segmental phenomena seems to involve the relationship between substanceless and substanceful nuclei and their onsets.[16] This relation cannot be given justice as long as no relation of licensing can be recognized in contexts other than Coda Mirror. It should be noted that most of the above problems are connected with static phonotactics and what we call indirect effects of government and licensing, that is, cases where, on top of the formally defined positional strength, a separate systemic decision is required whether a particular weakening or strengthening phenomenon is to be effected. Otherwise, it is rather encouraging that such a wide range of prosodic phenomena as those discussed in this text can be handled by such a small number of assumptions concerning the phonological representation above the melody, namely, that the skeleton is reduced to sequences of CVs, and that the entire job of organization of these positions is done by just two types of lateral relations: government and licensing, where the latter might need to be broken down into two types: a-licensing and p-licensing.

10.3 Further reading

Kaye, Jonathan 1990. 'Coda' licensing. *Phonology* 7: 301–330.
The central paper for constituent structure in Standard GP: this is where the idea comes from that word-final consonants are always the onset of an empty nucleus, in all languages.

Harris, John 1994. *English Sound Structure*. Oxford: Blackwell.
Summary of Standard GP: syllable structure (chapters 2 and 4.6), segmental structure (chapter 3), Licensing (chapter 4).

Scheer, Tobias 2004. *A Lateral Theory of Phonology: Vol.1: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.
§165 contains a description of the typology of lateral relations in Standard GP, critique thereof and transition to Strict CV where only two lateral relations are left which have stable effects (something that was not the case before): government and licensing. This book also proposes a short guide to Standard GP (§623).

Pöchtrager, Markus Alexander, and Jonathan Kaye 2013. GP2.0. *SOAS Working Papers in Linguistics and Phonetics* 16: 51–64.
The article exposes the workings of GP 2.0, where constituency and lateral relations are quite different from other versions of GP.

Lowenstamm, Jean 1996. CV as the only syllable type. *Current Trends in Phonology: Models and Methods*, Vol. 2, edited by Jacques Durand and Bernard Laks, 419–441. Salford, Manchester: ESRI.
The paper contains the original proposal that CV is the only syllable type and provides several arguments based on templatic and non-templatic languages.

Ségéral, Philippe, and Tobias Scheer 2001. La Coda-Mirroir. *Bulletin de la Société de Linguistique de Paris* 96: 107–152.
Offers one of the most important theoretical advances on the treatment of the strong and weak positions in GP, deriving causality of lenition and fortition from the arrangement of government and licensing in a Strict CV phonological representation.

Scheer, Tobias, and Marketa Ziková 2010. The Coda Mirror v2. *Acta Linguistica Hungarica* 57.4: 411–431.
Provides a revised version of Coda Mirror, which also includes the behaviour of branching onsets.

Cyran, Eugeniusz 2010. *Complexity Scales and Licensing in Phonology*. Berlin: Mouton de Gruyter.
The book introduces an alternative version of a CV model which incorporates the main aspects of Standard GP, including especially the concept of Government Licensing and scales relating both to formal complexity of syllabic configurations and relative strength of licensers.

Notes

1 Of course the detail is more intricate: individual languages make sovereign decisions as to what exactly is a good branching onset among those clusters that have a rising sonority profile (see Clements 1990). And of course parametric decisions control syllabification: some languages (Semitic for example) do not provide for branching onsets at all, and hence VTRV and VRTV will both end up with a heterosyllabic cluster (VT.RV and VR.TV).
2 Note that here and henceforth the symbol "ø" refers to an empty nucleus (rather than to a front rounded vowel).
3 Harris (1994: 184ff.) discusses similar evidence against resyllabification from English syncope, which is possible in *sep(a)rate* and *fam(i)ly*, but not in **ball(o)tting*.
4 Note that the same pattern holds for the two other mid vowels: *b*[e]*tise* 'silliness', *m*[e]*trique* 'metrical' vs. *al*[ɛ]*rter* 'to alarm', *b*[ɛ]*t(e)* 'silly'; *gu*[ø]*ler* 'to shout', *p*[ø]*pler* 'to populate' vs. *s*[œ]*rfer* 'to surf', *gu*[œ]*l(e)* 'snout'.
5 Positional strength expressed in syllabic terms by referring to such constructs as onset (strong) vs. coda (weak) is inadequate. Intervocalic onsets are notoriously weak and subject to lenition. A more precise model will be described in section 10.2.4 below.

6 In fact, there are a number of CV models entertained currently in which the major differences lie precisely in how the two lateral forces of government and licensing are utilized. The reader is referred to Bendjaballah (1999); Szigetvári (1999); Rowicka (1999); Kula & Marten (2009); Cyran (2010); Zdziebko (2012); Scheer (2004, 2012a) for more information.
7 For a wealth of argumentation and wide empirical coverage of LTP, as well as the diachronic development of this model, the reader is advised to read Scheer (2004) before Scheer (2012b).
8 There is a huge body of literature concerning vowel-zero alternations within Standard GP and Strict CV which is why this aspect of lateral relations is omitted here. The reader is referred to Kay et al. (1990); Charette (1990); Gussmann & Kaye (1993) for the former, and Rowicka (1999); Scheer (2004); Cyran (2010) for different takes within the latter. For more recent developments within LTP, see Scheer (2012c).
9 There is an additional complication which is discussed, for example, in Zdziebko (2012). Licensing of vowel length seems to be conditioned by the nature of the intervening consonant if the licenser is an empty nucleus (FEN), and to some extent even if the licenser is a full vowel. See also Bednarska (2015) for similar phenomena in Welsh and Breton.
10 We leave aside a third option, where the mono-positional behaviour gets a mono-positional representation (e.g. Lowenstamm 2003a; Rennison 1998). For a full discussion of this typology see Ségéral & Scheer (2005).
11 Details concerning the conditioning of this relation remain to be worked out. Clearly what is crucial is the nature of the two consonants in the TR sequence, and their melodic adjacency. There is also some conditioning connected with the nature of the following nucleus, in that some languages allow TRs to occur only before vowels, not before empty nuclei. This suggests that a relation of licensing is involved, but it is still unclear how. There are existing proposals concerning licensing of governing relations within Standard GP (Charette 1990), as well as within later models using the CV skeleton (e.g., Cyran 2010), which might be taken into account.
12 Broselow (2003) is an interesting exception.
13 This example really refers to the behaviour of the final C in Yawelmani as a coda, in that it closes the syllable and does not allow for a preceding long vowel. Of course, as with anything in this model, this follows from the activity or inactivity of the following nucleus, in this case the FEN.
14 A similar problem is found in the allophonic distribution of [w] and [v] in Belorussian (Scheer 2012a), where [w] seems to occur before empty nuclei, e.g. [korowka] 'cow, dim', [korow] 'cow, gen.pl.', and [v] in front of a vowel, e.g. [korova] 'cow, no.sg.', [vada] 'water', [barva] 'colouring'. Note that the 'stronger' [v] occurs in the weakest, intervocalic, context. Thus it is not positional strength that governs the distribution but rather the nature of the following nucleus (for more details see Chapter 11).
15 That IENs must be licensers also follows from a recent analysis of licensing of laryngeal specifications in Polish (e.g. Cyran 2014a), in which both true and bogus TRs retain the laryngeal category on the obstruent.
16 There is a substantial body of evidence showing that nuclei license their onsets as well as the preceding nucleus at the same time, and that the two 'loyalties' are mutually dependent. See, for example, Kula & Marten (2009); Zdziebko (2012); Cyran (2014a), and references therein.

References

Anderson, Stephen 1982. The analysis of French shwa: Or, how to get something for nothing. *Language* 58: 534–573.

Angoujard, Jean-Pierre 1982. Sur la représentation du verbe [ktib] (parler de Tunis). *Analyses, Théorie* 2–3: 1–17.

Barillot, Xavier and Philippe Ségéral 2005. On phonological Processes in the '3rd' conjugation in Somali. *Folia Orientalia* 41: 115–131.

Bednarska, Katarzyna 2015. Quantity in Breton and Welsh. Ph.D. dissertation, KUL Lublin.

Bendjaballah, Sabrina 1999. Trois figures de la structure interne des gabarits. Ph.D. dissertation, Université Paris 7.

Bendjaballah, Sabrina 2001. The negative preterite in Kabyle Berber. *Folia Linguistica* 34: 185–223.

Bendjaballah, Sabrina, and Martin Haiden 2008. A typology of emptiness in templates. *Sounds of Silence: Empty Elements in Syntax and Phonology*, edited by Jutta Hartmann, Veronika Hegedüs and Henk van Riemsdijk, 23–59. Amsterdam: Elsevier.

Brockhaus, Wiebke 1995. Skeletal and suprasegmental structure within government phonology. *Frontiers of Phonology*, edited by Jacques Durand and Francis Katamba, 180–221. London & New York: Longman.

Broselow, Ellen 2003. Marginal phonology: Phonotactics on the edge. *The Linguistic Review* 20: 159–193.

Brun-Trigaud, Guylaine, and Tobias Scheer 2010. Lenition in branching onsets in French and in ALF dialects. *Development of Language through the Lens of Formal Linguistics*, edited by Petr Karlík, 15–28. Munich: Lincom.

Burzio, Luigi 1994. *Principles of English Stress*. Cambridge: Cambridge University Press.

Caratini, Emilie 2009. Vocalic and consonantal quantity in German: Synchronic and diachronic perspectives. Ph.D. dissertation, Nice University and Leipzig University.

Carvalho, Joaquim Brandão de 2002. De la syllabation en termes de contours CV. Habilitation Thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris. WEB.

Carvalho, Joaquim Brandão de 2004. Templatic morphology in the Portuguese verb. *Nouveaux départs en phonologie: les conceptions sub- et suprasegmentales*, edited by Trudel Meisenburg and Maria Selig, 13–32. Tübingen: Gunter Narr.

Carvalho, Joaquim Brandão de 2008. From positions to transitions: A contour-based account of lenition. *Lenition and Fortition*, edited by Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 415–445. Berlin: de Gruyter.

Charette, Monik 1990. Licence to govern. *Phonology* 7: 233–253.

Charette, Monik 2008. The vital role of the trochaic foot in explaining Turkish word endings. *Lingua* 118: 46–65.

Clements, George 1990. The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology I*, edited by John Kingston and Mary Beckmann, 283–333. Cambridge: Cambridge University Press.

Cyran, Eugeniusz 2010. *Complexity Scales and Licensing in Phonology*. Berlin: Mouton de Gruyter.

Cyran, Eugeniusz 2014a. *Between Phonology and Phonetics: Polish Voicing*. Berlin and New York: De Gruyter Mouton.

Cyran, Eugeniusz 2014b. The phonology and phonetics of obstruentization. *Crossing Phonetics-Phonology Lines*, edited by Eugeniusz Cyran and Jolanta Szpyra-Kozłowska, 31–46. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Cyran, Eugeniusz, and Edmund Gussmann 1999. Consonantal clusters and governing relations: Polish initial consonant sequences. *The Syllable. Views and Facts*, edited by Harry van der Hulst and Nancy Ritter, 219–247. Berlin: Mouton de Gruyter.

Dell, François 1995. Consonant clusters and phonological syllables in French. *Lingua* 95: 5–26.

Durand, Jacques 1990. *Generative and Non-Linear Phonology*. London & New York: Longman.

Emonds, Joseph 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.

Faust, Noam 2014. Templatic metathesis in Tigre imperatives. *Phonology* 31: 209–227.

Faust, Noam 2015. Eroded prefixes, gemination and guttural effects in Tigre: An account in CVCV phonology. *Natural Language and Linguistic Theory* 33: 1209–1234.

Fortuna, Marcin 2015. *Double Licensing Phonology*. München: Lincom.

Fortuna, Marcin 2016. Icelandic post-lexical syllabification and vowel length in CVCV phonology. *The Linguistic Review* 33: 239–275.

Gussmann, Edmund 2002. *Phonology: Analysis and Theory*. Cambridge: Cambridge University Press.

Gussmann, Edmund 2007. *The Phonology of Polish*. Oxford: Oxford University Press.

Gussmann, Edmund, and Jonathan Kaye 1993. Polish notes from a Dubrovnik Café: I. The yers. *SOAS Working Papers in Linguistics and Phonetics* 3: 427–462.

Harris, John 1990. Segmental complexity and phonological government. *Phonology* 7: 255–300.

Harris, John 1994. *English Sound Structure*. Oxford: Blackwell.

Harris, John 1997. Licensing Inheritance: An integrated theory of neutralisation. *Phonology* 14: 315–370.
Hayes, Bruce, Robert Kirchner, and Donca Steriade (eds.) 2004. *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.
Kaye, Jonathan 1990. '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. Reprinted in *A Festschrift for Edmund Gussmann*, edited by Henryk Kardela and Bogdan Szymanek, 155–176. Lublin 1996: Lublin University Press.
Kaye, Jonathan 1995. Derivations and interfaces. *Frontiers of Phonology*, edited by Jacques Durand and Francis Katamba, 289–332. London & New York: Longman. Also in *SOAS Working Papers in Linguistics and Phonetics* 3, 1993, 90–126.
Kaye, Jonathan, and Jean Lowenstamm 1981. Syllable structure and markedness theory. *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, edited by Adriana Belletti, Luciana Brandi and Luigi Rizzi, 287–315. Pisa: Scuola normale superiore.
Kaye, Jonathan, and Jean Lowenstamm 1984. De la syllabicité. *Forme Sonore du Langage*, edited by François Dell, Daniel Hirst and Jean-Roger Vergnaud, 123–159. Paris: Hermann.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. *Phonology* 7: 193–231.
Kenstowicz, Michael 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.
Kiparsky, Paul 1991. *Catalexis*. Ms., Stanford University.
Kula, Nancy, and Lutz Marten 2009. Defining initial strength in clusterless languages in Strict CV. *Strength relations in phonology*, edited by Kuniya Nasukawa and Phillip Backley, 251–284. Berlin: Mouton de Gruyter.
Lahrouchi, Mohamed 2003. Manifestations gabaritiques dans la morphologie verbale du berbère tachelhit. *Recherches Linguistiques de Vincennes* 32: 61–82.
Lahrouchi, Mohamed 2008. A templatic approach to gemination in the imperfective stem of Tashlhiyt Berber. *Studies in African Linguistics* 37: 21–60.
Lahrouchi, Mohamed, and Philippe Ségéral 2009. Morphologie gabaritique et apophonie dans un langage secret féminin en berbère tachelhit. *Revue Canadienne de Linguistique* 54: 291–316.
Lowenstamm, Jean 1996. CV as the only syllable type. *Current Trends in Phonology: Models and Methods*, vol. 2, edited by Jacques Durand and Bernard Laks, 419–441. Salford, Manchester: ESRI.
Lowenstamm, Jean 1999. The beginning of the word. *Phonologica 1996*, edited by John Rennison and Klaus Kühnhammer, 153–166. La Hague: Holland Academic Graphics.
Lowenstamm, Jean 2000a. The image of a segment. *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*, edited by Chris Schaner-Wolles, John Rennison and Friedrich Neubarth, 281–290. Torino: Rosenberg & Sellier. WEB.
Lowenstamm, Jean 2000b. The no straddling effect and its interpretation: A formal property of chaha 2nd feminine singular formation. *Research in Afroasiatic Grammar*, edited by Jacqueline Lecarme, Jean Lowenstamm and Ur Shlonsky, 183–198. Amsterdam: Benjamins.
Lowenstamm, Jean 2003a. Remarks on mutae cum liquida and branching onsets. *Living on the Edge: 28 Papers in Honour of Jonathan Kaye*, edited by Stefan Ploch, 339–363. Berlin: Mouton de Gruyter.
Lowenstamm, Jean (ed.) 2003b. *Grammaire et gabarits*. St.-Denis: Recherches Linguistiques de Vincennes 32.
Oostendorp, Marc van 2005. The first person singular in Dutch dialects. *Proceedings of the Thirty-Fifth Annual Meeting of the North East Linguistic Society*, edited by Leah Bateman and Cherlon Ussery, 1–12. Amherst, MA: GLSA.
Passino, Diana 2009. The allomorphy of nen as a case of vowel/zero alternation in Teraman Abruzzese. *Studi e Saggi Linguistici* 47: 83–110.
Passino, Diana 2013. A unified account of consonant gemination in external sandhi in Italian: Raddoppiamento Sintattico and related phenomena. *The Linguistic Review* 30: 313–346.
Piggott, Glyn 1999. At the right edge of words. *The Linguistic Review* 16: 143–185.

Pöchtrager, Markus Alexander 2006. The structure of length. Ph.D. dissertation, University of Vienna. WEB.

Pöchtrager, Markus Alexander, and Jonathan Kaye 2013. GP2.0. *SOAS Working Papers in Linguistics and Phonetics* 16: 51–64.

Polgárdi, Krisztina 1998. *Vowel Harmony: An Account in Terms of Government and Optimality*. The Hague: Holland Academic Graphics.

Polgárdi, Krisztina 1999. Constraint ranking, government licensing and the fate of final empty nuclei. *Phonologica 1996*, edited by John Rennison and Klaus Kühnhammer, 167–182. The Hague: Holland Academic Graphics.

Polgárdi, Krisztina 2003. Hungarian as a strict CV language. *The Phonological Spectrum: Vol II: Suprasegmental Structure*, edited by Harry van der Hulst, Vincent van Heuven and Jeroen van de Weijer, 59–79. Amsterdam & Philadelphia: Benjamins.

Polgárdi, Krisztina 2008. The representation of lax vowels in Dutch: A loose CV approach. *Lingua* 118: 1375–1392.

Polgárdi, Krisztina 2009. Trochaic proper government, loose CV, and vowel-zero alternation in Hungarian. *Approaches to Hungarian, Vol.11: Papers from the 2007 New York Conference*, edited by Marcel den Dikken and Robert Vago, 143–165. Amsterdam: Benjamins.

Rennison, John 1998. Contour segments without subsegmental structures. *Structure and interpretation. Studies in phonology*, edited by Eugeniusz Cyran, 227–245. Lublin: Folium.

Rizzolo, Olivier 2007. The syllable is not a valid constituent: Evidence from two Serbo-Croatian language games. *Formal Approaches to Slavic Linguistics 15: The Toronto Meeting*, edited by Richard Compton, Magda Goledzinowska and Ulyana Savchenko, 264–281. Ann Arbor: Michigan Slavic Publications.

Rowicka, Grażyna 1999. *On Ghost Vowels: A Strict CV Approach*. The Hague: Holland Academic Graphics.

Rubach, Jerzy 1999. The syllable in phonological analysis. *Rivista di Linguistica* 11: 273–314.

Scheer, Tobias 1996. Une théorie de l'interaction directe entre consonnes. Ph.D. dissertation. Paris: Universiteé Paris 7.

Scheer, Tobias 2004. *A Lateral Theory of Phonology: Vol.1: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.

Scheer, Tobias 2011. *A Guide to Morphosyntax-Phonology Interface Theories: How Extra-Phonological Information Is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.

Scheer, Tobias 2012a. At the right edge of words (again). *McGill Working Papers in Linguistics* 22: 1–29. Available at www.mcgill.ca/mcgwpl/archives/volume-221-2012.

Scheer, Tobias 2012b. *Direct Interface and One-Chanel Translation: Vol. 2 of a Lateral Theory of Phonology*. Berlin: Mouton de Gruyter.

Scheer, Tobias 2012c. Variation is in the lexicon: Yer-based and epenthetic vowel-zero alternations in Polish. *Sound, Structure and Sense: Studies in Memory of Edmund Gussmann*, edited by Eugeniusz Cyran, Henryk Kardela and Bogdan Szymanek, 632–671. Lublin: Wydawnictwo KUL.

Scheer, Tobias 2015. *Précis de structure syllabique. Accompagné d'un apparat critique*. Lyon: ENS Editions.

Scheer, Tobias, and Philippe Ségéral 2008a. Positional factors in lenition and fortition. *Lenition and Fortition*, edited by Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 131–172. Berlin: Mouton de Gruyter.

Scheer, Tobias, and Philippe Ségéral 2008b. The coda mirror, stress and positional parameters. *Lenition and Fortition*, edited by Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 483–518. Berlin: Mouton de Gruyter.

Scheer, Tobias, and Marketa Ziková 2010. The coda mirror v2. *Acta Linguistica Hungarica* 57.4: 411–431.

Ségéral, Philippe, and Tobias Scheer 2001. La Coda-Mirroir. *Bulletin de la Société de Linguistique de Paris* 96: 107–152.

Ségéral, Philippe, and Tobias Scheer 2005. What lenition and fortition tells us about Gallo-Romance *Muta cum Liquida*. *Romance Languages and Linguistic Theory 2003*, edited by Twan Geerts, Ivo van Ginneken and Haike Jacobs, 235–267. Amsterdam: Benjamins.
Sievers, Eduard 1885. *Grundzüge der Phonetik*. Vol. 3. Auflage Leipzig: Breitkopf and Härtel.
Sievers, Eduard 1901. *Grundzüge der Phonetik: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Vol. 5. Auflage Leipzig: Breitkopf & Härtel.
Spencer, Andrew 1986. A non-linear analysis of vowel-zero alternations in Polish. *Journal of Linguistics* 22: 249–280. WEB.
Steriade, Donca 1982. Greek Prosodies and the Nature of Syllabification. Ph.D. dissertation, MIT.
Steriade, Donca 1999. Alternatives to syllable-based accounts of consonantal phonotactics. *Proceedings of LP'98: Item Order in Language and Speech*, Vol. 1, edited by Osamu Fujimura, Brian D. Joseph and Bohumil Palek, 205–245. Prague: Karolinum Press.
Szigetvári, Péter 1999. Why CVCV? *The Even Yearbook* 4: 117–152.
Szigetvári, Péter 2001. Dismantling syllable structure. *Acta Linguistica Hungarica* 48: 155–181.
Szigetvári, Péter 2007. Branching onsets and syncope in English. *Language Sciences* 29: 408–425.
Szigetvári, Péter 2008. Two directions for Lenition. *Lenition and Fortition*, edited by Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 561–592. Berlin: Mouton de Gruyter.
Takahashi, Toyomi 1993. A farewell to constituency. *UCL Working Papers in Linguistics* 5: 375–410.
Ulfsbjorninn, Shanti 2014. A field theory of stress: The role of empty nuclei in stress systems. Ph.D. thesis. London: SOAS.
Yoshida, Shohei 1993. Licensing of empty Nuclei: The case of Palestinian vowel harmony. *The Linguistic Review* 10: 127–159. WEB.
Zdziebko, Sławomir 2012. *Issues in Scottish Vowel Quantity*. Newcastle upon Tyne: Cambridge Scholars Publishing.
Zdziebko, Sławomir 2015. Licensing of vowel quantity and the Scottish Vowel Length Rule. *Roczniki humanistyczne* 63: 151–176.

11
Interfaces in Government Phonology

Tobias Scheer and Eugeniusz Cyran

11.1 Interface with morpho-syntax

11.1.1 General (and pre-theoretical) settings

11.1.1.1 Distinct computational systems and their communication

Before describing the specific contribution of Government Phonology (GP) to the workings of the interface with morpho-syntax, it is useful to recall a number of general issues that all interface theories need to address. The broadest of these is the modular architecture of grammar that lies at the heart of the generative enterprise since *Aspects* (Chomsky 1965; see Chapter 9.2.1). That is, grammar is made of a number of distinct computational systems, each of which operates over a proprietary vocabulary that is distinct from the ones used by other systems. Taking these vocabulary items that are stored in long-term memory as an input, computational systems build structure (e.g. trees in syntax) following hard-wired instructions (among others, Merge in syntax) in online processing (active memory). The output may then be communicated to other modules for further processing, but the distinct vocabulary sets (called *domain specificity* in cognitive science) blocks direct transmission of information. Hence the output of a donor module needs to be translated into the idiom of the receiving module prior to transmission. This is what we call the interface, or interface operations: making information legible for the recipient.[1]

For the purpose of this chapter, it is enough to agree that morpho-syntax and phonology are distinct computational systems, one operating over lexical items such as number, person, animacy etc., the other working with labiality, occlusion and the like (no overlap). A further point is the fact that in production morpho-syntax feeds phonology in the (functional) sense that it concatenates items retrieved from long-term memory; the product of this gluing-together is linearized and enters phonological computation as a linear string.[2] Phonology itself does not concatenate anything, nor create linearity: it interprets whatever is delivered. Hence morpho-syntactic activity necessarily occurs prior to the workings of phonology (in production).

11.1.1.2 Morphological vs. syntactic influence on phonology

Another general setting that holds for the entire interface literature is this: what is studied almost exclusively is the influence of morpho-syntax on phonology, not the reverse. Zwicky & Pullum (1986) have introduced the idea that syntax is phonology-free, i.e. uninfluenced by any phonological information. This principle was challenged namely on the grounds of intonation, stress, rhythm, minimal word constraints and other properties that are located at or above the skeleton in phonological representations. Melody (i.e. items below the skeleton), however, never appears to bear on syntax. The correct generalization thus seems to be melody-free syntax (Scheer 2011: §662, 2016a). In any event, this chapter, following the literature, is only concerned with the influence of morpho-syntax on phonology (not the reverse).

Another question concerns an eventual distinction between morphology and syntax. This is a long-standing debate in itself: traditional generative and non-generative approaches (today autonomous morphology, i.e. the heirs of Lexical Phonology/Morphology) devise two distinct systems, but the trend in generative quarters is to consider that morphology and morphology follow the same workings and hence constitute a single computational system. The latter view is held by Distributed Morphology (e.g. Embick & Noyer 2007) and Nanosyntax (e.g. Starke 2009). Interface theories (which are almost always made by phonologists) clearly reproduce the distinction between syntax and morphology (see Scheer 2011: §423). Roughly, Lexical Phonology is responsible for morphological influence (i.e. by smaller pieces below the word level, with a cyclic/derivational management), while Prosodic Phonology describes how syntax bears on phonology (i.e. larger pieces at and above the word size: postlexical phonology, with a representational management).

Until the late 1990s, the little that was done in GP regarding the interface with morpho-syntax (domain structure) concerned only morphology, and hence had Lexical Phonology as a reference point. This will be explained in some detail below. External sandhi, i.e. the bearing of syntactic divisions on phonology at and above the word level, only entered the scene when the so-called initial CV was introduced (Lowenstamm 1999) and combined with syntactic phase theory (Chomsky 2000; on which more in sections 11.1.2.3 and 11.1.2.4 below).

11.1.1.3 Derivational vs. representational management of the interface

The final point to be made in this introduction is Interface Dualism (Scheer 2011: §6). There are two ways for morpho-syntax to influence phonology: derivationally and representationally. The former is a genuinely generative invention that came into being in Chomsky et al. (1956: 75) and was successively known as the transformational cycle, the phonological cycle, cyclic derivation and finally today derivation by phase (in syntactic quarters). It embodies the insight that (phonological and semantic) interpretation applies successively from the most to the least embedded piece of structure.

The other means by which morpho-syntax can bear on phonology is through the insertion of a representational object into the linear string that is submitted to phonological computation. This is the traditional interface management which is practiced (at least) since the 19th century, and in any case is shared by structuralist and generative thinking: carriers of extra-phonological information in phonology have been successively incarnated as juncture phonemes, SPE-type diacritics (# and the like) and the Prosodic Hierarchy (ω, φ and so forth), each being the representative of its time, i.e. reflecting general assumptions on the organization of phonological units (phonemes, segments, autosegmental structure).

This is also the division of labour that underlies interface thinking. It was mentioned that Lexical Phonology (which proposes a purely derivational management) is supposed to account for pieces up to the word level, i.e. within the realm of morphology, and explicitly rejects any procedural management of external sandhi (postlexical phonology is supposed to be non-cyclic). Prosodic Phonology, on the other hand, uses only representational devices, i.e. constituents of the Prosodic Hierarchy, whose higher levels (from the Prosodic Word on) describe the influence of pieces that have word size or are bigger (syntax).

In this context, GP was only derivational in the 1990s, rejecting any representational management of the interface (Kaye 1995). This is not unrelated to the aforementioned fact that the reference point for Kaye was Lexical Phonology. Here again, the aforementioned initial CV modifies the picture substantially: it introduces a representational carrier of morpho-syntactic information in phonology (see section 11.1.3.1).

11.1.2 Standard GP: domain structure

11.1.2.1 Visible and invisible morpheme boundaries

In Standard GP, i.e. up to the late 1990s, there were basically two articles about how GP views the interface with morpho-syntax: Kaye (1992) and Kaye (1995), plus relevant pieces of Prunet's (1986) dissertation (including Prunet 1987). A more complete overview of Kaye's interface theory is discussed in Scheer (2011: §258).

Kaye's basic idea is that a morpheme boundary may either be visible or invisible to phonological computation. In Kaye's vocabulary, the former is called analytic morphology (the phonological string is analyzed into two substrings), while the latter amounts to non-analytic morphology (the phonological string, although morphologically complex, is phonologically unanalyzable). That is, a morphological structure [[A]B] (where A and B are morphemes) appears as [AB] in phonology: the fact that B is a separate morpheme goes unnoticed (the boundary is not flagged), and hence the computation of [AB] is indistinguishable from the computation of a monomorphemic string. For instance, stress in English (which for the sake of exposition we will assume to be simply penultimate) falls on the second but last vowel in monomorphemic items (*párent*) as much as in morphologically complex strings made of a root and a class 1 suffix (*parént-al*). Kaye concludes that class 1 affixes are of the invisible kind: stress assignment operates over [parent-al] as much as it does over [parent].

Class 2 affixes, on the other, hand are visible in phonology: they are flagged and leave a trace. Hence a morphological structure [[A]B] is computed as such in the phonology. Chunks that are submitted to phonological interpretation and computed in one go (here [A] and [AB]) are called domains in Kaye's vocabulary – they are known as cycles (or levels) and today as phases.

Phonologically visible morpheme boundaries alter regular phonological computation in the following way: properties of the string that are acquired by computation in a given domain cannot be modified by further computation occurring in outer domains (Kaye calls this robustness). Hence when a class 2 suffix such as *-hood* attaches to a root as in [[parent]-hood], stress is first assigned to the inner domain, producing *párent*. Stress assignment is then also performed on the outer domain [párent-hood], but given robustness the previously acquired location of stress cannot be undone or modified, and the result is *párent-hood*.

With the notion of robustness, Kaye in fact pioneers what today is known as the Phase Impenetrability Condition (PIC) in (syntactic) phase theory (on which more in section 11.1.2.3). Another example of the prohibition to undo what was done in previous domains comes from French and is exposed both in Kaye (1992: 142ff.) and Kaye (1995: 306ff.). Data and analysis are originally due to Prunet (1986, 1987). French shows a contrast between *mon ami* [mɔ̃n ami] 'my friend' and *bon ami* [bɔn ami] 'good friend': the vowel of the possessive (*mon* 'my') is nasalized, while the vowel of the adjective (*bon* 'good') is not.

Both determiners bear a liaison consonant at their right edge: the *-n* is only present when the following noun begins with a vowel. It is absent before consonants (*mon café* [mɔ̃ kafe] 'my coffee', *bon café* [bɔ̃ kafe] 'good coffee') and when the words are pronounced in isolation (*mon* [mɔ̃] 'my', *bon* [bɔ̃] 'good'). Note that in these contexts the vowel of both *mon* and *bon* is consistently nasalized. Following standard autosegmental assumptions on French liaison, liaison consonants are lexically floating.

Prunet and Kaye argue that there is good reason to believe that the syntactic relationship of *bon* is closer than the one that *mon* entertains with the following noun. This syntactic difference is then transmitted to the phonology in terms of domain structure. While *mon ami* is the complex [[mon] ami], *bon ami* lacks internal structure: it identifies as [bon ami].

(1) French *mon ami* vs. *bon ami*: input to the phonology

```
a. mon ami                          b. bon ami
   O N       O N O N                   O N    O N O N
   | |       | | | |                   | |    | | | |
[[ x x    ] x x x x ]               [ x x    x x x x ]
   | |         | | |                   | |      | | |
   m o n       a m i                   b o n    a m i
```

Two processes now apply: floating consonants associate to available consonantal positions (the floating nasal here behaves like all other floating consonants), and nasalization of vowels is effected by nasals that occur domain-finally or before a consonant (this is a specific behavior of nasals, but general in the language). The derivation proceeds cyclically: phonology is first done on the inner domain of [[mon] ami]; the result is vowel nasalization, i.e. the association of the floating nasal to the preceding nucleus. On the outer cycle, liaison also associates the nasal to the following Onset, which is now available. In the end, the nasal consonant thus enjoys a double association: it contributes to the pronunciation of the preceding nucleus and the following Onset; the result is [mɔ̃n ami]. By contrast, the nasal in [bon ami] will undergo liaison, but fails to nasalize the preceding vowel because *bon* is not a domain by itself – hence the nasal is never domain-final. Therefore the triggering environment for nasalization is never met, and the result is [bɔn ami].

Critical for this analysis is that the association of the nasal to the preceding nucleus that is achieved on the inner domain is not undone when the outer domain is computed. This is due to robustness (Kaye's PIC).

11.1.2.2 Parsing cues

Kaye (1989, 1995) follows the perception-oriented logic of Trubetzkoy's *Grenzsignale* (which has a number of modern incarnations; see for example Hume & Johnson 2001; Boersma 1998). The idea is that phonology helps the listener to identify morphemes in the unstructured linear signal: this signal contains information, parsing cues, which flag morpheme boundaries. When the phonological system of the listener runs over the linear string, it is able to tell whether or not a given sequence is well-formed according to its standards: a sequence is ill-formed if it does not conform to what the application of phonological computation would have produced. In other words, speakers know what a morpheme may look like, and what it cannot look like. Anything that constitutes a phonological anomaly or is not compatible with morpheme structure, then, sends an alarm to the parsing system: it signals the presence of a morpheme boundary.

The cluster [mz], for example, is a possible sequence in English (*it seem-s*, *dream-s* etc.), but does not occur within a morpheme. The same is true for the cluster [stm] that is found in *postman*. In these cases, a monomorphemic parse of the string is thus incompatible with the phonological knowledge of the listener. Hence the parses [[seem] s] and [[post] man] are enforced. Another example is the sequence [ksθs] in the word *sixths*. Since neither [ksθ] nor [sθs] nor [sθ] nor [θs] can be tautomorphemic, the correct structure must be [[[ksø]θø]sø]. In the same way, *darkness* and *enlargement* could not be the result of a computation over a single domain either since [rkn] and [rd͡ʒm] are not good tautomorphemic clusters. This is how the suffixes *-ness* and *-ment* are identified by phonology alone, i.e. without look-up in the lexicon or any morphology intervening.

While the cases of morpheme-detection described are theory-independent, there are also theory-specific parsing cues. For example, Standard GP holds that super-heavy rhymes (a long vowel plus a coda, i.e. a rhyme dominating more than two skeletal slots) are universally ill-formed (a consequence of the Binary Theorem; see Kaye 1990 and section 1.3 of Chapter 10). The consonant of VVC sequences is therefore always an Onset, even when it is word-final or followed by another consonant. Hence speakers know that the only possible parse of VVCC# sequences contains two empty nuclei, /VV.Cø.Cø#/. Word-final consonants are onsets of empty nuclei anyway, and the cluster-internal empty nucleus is the only possible parse given the ban on super-heavy rhymes. In other words, a universal property of syllable structure betrays the existence of empty nuclei in words such as *seem-ed* [siimd], *peep-ed* [piipt], *seep-ed* [siipt] and *fak-ed* [fejkt], whose only possible parses are /[[siimø]dø]/, /[[piipø]dø]/, /[[siipø]dø]/ and /[[fejkø]dø]/, respectively. In each case, the long vowel (or heavy diphthong) allows only for an Onset interpretation of the following consonant, which in turn can only be followed by another Onset consonant.

11.1.2.3 Location in the general landscape: Lexical Phonology, phase theory

Let us now look at the formal characteristics of Kaye's domains. In a given language, Kaye calls the set of phonological instructions that effect phonological computation the φ-function. The internal workings of the φ-function are described in section 1.4.2 of Chapter 9 (no extrinsic ordering or ranking of instructions). An important property of the φ-function for the present purpose is that there is only one. Almost all phonological theories implement a number of different phonologies, i.e. distinct computational systems that apply phonological instructions in different ways. Mini-phonologies in a given language may be chunk-specific (i.e. specific to a particular size of the input string: general vs. word-level phonology in SPE, stem-level vs. word-level vs. postlexical phonology in Lexical Phonology and Stratal OT) or morpheme-specific (i.e. specific to a particular class of morphemes: indexed constraints, cophonologies). According to Kaye, phonology is one: there is only one φ-function per language.[3]

Technically speaking, the φ-function:

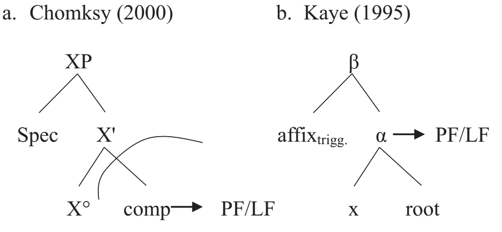
> has one argument, a phonological string, and returns the application of the phonology to this argument, also a phonological string. The expression φ(X) means, "apply phonology to the string X". φ(X) returns the phonological string which results from the application of phonology to its argument.
>
> *(Kaye 1995: 302)*

Kaye builds on interactionism, the central insight of Lexical Phonology. That is, morphological (concatenation) and phonological (interpretation) activity is intertwined: first you do phonology on an item, then you concatenate another piece, then you do phonology again on the result of concatenation, and so forth. In Kaye's system, morphological activity is represented by the concat function (which today would be called Merge): φ(concat(X,Y)) means that phonology is done on the result of a morphologically complex string, but that this morphological complexity is invisible to the phonology (were [XY] monomorphemic, the result would be identical). Analytic (or cyclic) domain structure is created when concat and phonological interpretation are interleaved: given two morphemes X and Y, either may be subjected to φ before concatenation takes place. This situation corresponds to the expressions φ(concat(φ(X),Y)) and φ(concat(X,φ(Y))). In the former case, phonology operates over morpheme X, the result is concatenated with morpheme Y and phonology again applies to the output. The latter configuration is the symmetric counterpart.[4]

In SPE and GB syntax of the 1980s, all concatenation was completed before phonological and semantic interpretation started. Derivation by phase (Chomsky 2000 and following) abandons this scenario and adopts interactionism (which was introduced by Lexical Phonology). It also implements selective spell-out, i.e. the idea that not all morpheme breaks define a cycle (as was the case in SPE): only a subset of nodes of the syntactic structure, phase heads, trigger spell-out. Selective spell-out was introduced by Halle & Vergnaud (1987) in phonology (see Scheer 2011: §225 for more detail), and Kaye takes over this mechanism: the morpheme boundary in [parent-al] is invisible for phonology because it does not trigger spell-out. By contrast, the boundary in [[parent]-hood] is visible, which means that it has triggered spell-out: phonological computation applies to [parent] alone. In other words, class 1 affixes are interpretation-neutral, while class 2 affixes trigger interpretation, i.e. project phase heads.

This leads to another prominent property of current syntactic phase theory, which it turns out was anticipated by Kaye: the fact that when spell-out is triggered by a given piece, it is not this piece itself that is sent to interpretation, but its sister (see Scheer 2008a for more detail). In syntactic phase theory, the spell-out of an XP only triggers the interpretation of the complement, i.e. the sister of the head of the XP; the head and Spec, XP – called the phase edge – are spelt out only at the next higher phase (Chomsky 2000: 108). This allows material that is trapped in the spell-out domain to escape the PIC by moving to the edge, and hence to be available for further computation. The mechanism devised by Kaye also spells out the sister of the interpretation-triggering piece: *-hood* creates a domain that is subjected to interpretation, and this domain is its sister (i.e. excluding the head = the suffix itself). The parallel with syntax is depicted under (2) below where in both cases the sister of the lexical category that triggers spell-out (X°, $\text{affix}_{\text{trigg}}$) is actually spelt out: the complement and the node that dominates the root (α).

(2) The phase edge in syntax and phonology: spell-out your sister



Finally, let us look at the Phase Impenetrability Condition (PIC) as entertained in current syntactic phase theory. The PIC is a device which guarantees that previously interpreted strings do not burden further computation – in Chomsky's terms, strings that are returned from interpretation are 'frozen' and 'forgotten' when concatenation resumes. The history of no-look-back devices in generative theory starts with Chomsky's (1973) Conditions on Transformations, and its offspring – until its recent revival in the guise of the PIC – was essentially phonological. No-look-back devices are designed in order to prevent computation to consider 'old' strings. Depending on their precise formulation, however, they have quite different empirical effects, which correspond to the thing that the analyst wants the computation to be unable to do. Like for the phase edge, Kaye's modification-inhibiting robustness is a phonological precedent of Chomsky's PIC, though not the first implementation of this idea: the Free Element Condition (Prince 1985) restricts rules that erect foot structure to strings that do not possess any such structure yet. In other words, phonological computation can build but not destroy existing structure. The construction of syllable structure was restricted in the same way (e.g. van Oostendorp 1994). On the syntactic side, Riemsdijk's (1978: 160) Head Constraint is a precursor of modification-inhibiting no-look-back.[5]

In sum, Kaye's domain structure was clearly ahead of its time: it combines or introduces workings that today lie at the heart of current syntactic theory: interactionism, selective spell-out, the phase edge and the PIC.

11.1.3 A non-diacritic theory of the interface (Direct Interface)

11.1.3.1 The initial CV: TR-only vs. anything-goes languages

Let us now turn to the representational side of the interface. It was mentioned that GP developed interest in this aspect only when Strict CV entered the scene. Lowenstamm (1999) proposed that the phonological exponent of the morphological information 'beginning of the word' is an empty CV unit.[6] Among other things, his goal was to derive the typology of restrictions regarding word-initial clusters: some languages admit only TR clusters (English, Italian etc.), while others allow for any sonority slope (e.g. Moroccan Arabic). Let us call the former TR-only, the latter anything-goes languages (T stands of any obstruent, R for any sonorant). A third pattern is trivial, i.e. languages that do not allow for any word-initial clusters at all. What needs to be explained, then, is the absence of RT-only languages, i.e. cases where RT but not TR clusters may occur word-initially. Or, put differently, the implicational relationship between initial TR and RT clusters raises the question: why is it that a language which has the latter will also have the former, but one that has the former may or may not have the latter?

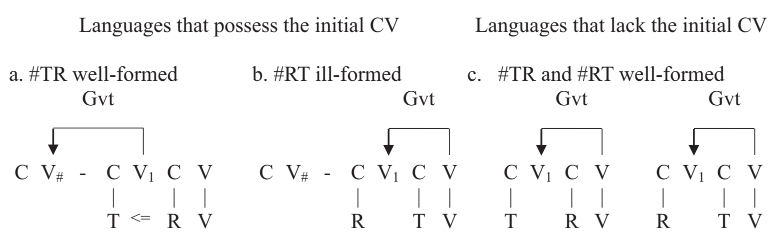
Lowenstamm's analysis is based on the identity of branching onsets (i.e. TR clusters) that was developed by Scheer (1999, 2004: §14): the solidarity of TR is due to a lateral relation among the two consonants that R is the head of (Infrasegmental Government, see section 2.3.4 of Chapter 10). This relation circumscribes the empty nucleus separating the two consonants. It could not relate the two consonants of an RT cluster since all lateral relations are head-final in Strict CV, and the head R could not be government-licensed by an empty nucleus. Table (3) shows how the initial CV ($CV_{\#}$) supplemented with the Strict CV analysis of branching onsets ('<=' represents the solidarity-creating lateral relation) derives the typological situation described.

(3b) is ill-formed because two empty nuclei occur in a row where the second (V_1) is unable to dispense government because it is itself governed. By contrast, under (3a) V_1

does not need to be governed by the following full vowel because it is enclosed within the solidarity domain of the TR cluster. Therefore it is a sound governor and may govern the empty nucleus of the initial CV. Hence the presence of the initial CV is responsible for the TR-only restriction: it adds a burden (an empty nucleus) that needs to be taken care of. Its absence in anything-goes languages takes away this burden, and accordingly there is only ever one empty nucleus that needs to be managed: the one enclosed within the initial cluster. As shown under (3c), this nucleus will always be able to be governed by the following full vowel, and hence the sonority slope of the surrounding consonants plays no role: any sonority sequence will do.

While for Lowenstamm (1999) the initial CV was always present but somehow switched off or made invisible in anything-goes languages, Scheer (2000: 274ff., 2004: §404) introduces the idea that its distribution is parameterized: languages decide whether or not to flag the left edge of the word just as they decide whether or not to spell out this or that syntactic property (person, case etc.) with a phonological marker of its own. The decision is parametric and hard-wired in the spell-out mechanism. But once it is taken, it is interpreted by the phonology, whose own workings produce the asymmetric picture at hand.

(3) Restrictions on initial clusters in CVCV



11.1.3.2 Non-arbitrary effects of the left edge of the word: three for the price of one

There are (at least) two other phenomena that are pervasively observed at the left edge of words and allow languages to pick one of two options. In some languages, first vowels of words cannot alternate with zero (while vowels elsewhere in the word can). In other languages first vowels do not show any peculiar behavior with respect to other vowels. But there is no language where non-initial vowels are unable to alternate with zero while initial vowels are. Relevant evidence is discussed in Ségéral & Scheer (2008) and Scheer (2004: §90). In strict CV, the reason why first vowels cannot alternate with zero is the presence of the initial CV: the relevant configuration is shown under (3b), only that the governed nucleus accommodates a vowel-zero alternation. Zero surfaces under government, which creates an ill-formed sequence of two empty nuclei. Therefore first vowels of words resist vowel-zero alternations in languages that mark the left edge of words with the initial CV. In absence of the initial CV, however, nothing withstands first vowels of words to alternate with zero.

The second phenomenon concerns word-initial consonants. In some languages, word-initial consonants are especially strong (and then pattern with post-coda consonants), while in others they show weak intervocalic behavior. But there is no language where they are especially weak. In terms of the Coda Mirror, a theory of lenition and fortition (see Chapter 10.2.3.1 and Chapter 10.2.3.2; Ségéral & Scheer 2008 for a summary), the strength of word-initial consonants is a consequence of the presence of an empty nucleus to their left:

the government of the following vowel is absorbed by the initial CV, which means that the consonant itself is licensed (i.e. backed up) but ungoverned (i.e. unspoiled), that is, experiences maximally comfortable conditions whose expression is segmental strength. In the absence of the initial CV, however, the vowel following initial consonants has no governing duties and therefore governs its own Onset – which is the description of regular intervocalic consonants (see Chapter 10.2.3.2).[7]

For each of the three phenomena discussed (restrictions on initial clusters, alternation of first vowels with zero and the strength of initial consonants), the cross-linguistic variation observed is due to the presence or absence of the initial CV. Since the presence of the initial CV is the result of a parametric choice, the prediction is made that any language which displays one of the three consequences of the initial CV (TR-only, first vowel unable to alternate with zero, initial consonant strong) will also instantiate the two others. And conversely, languages that display one of the three correlates of the absence of the initial CV (anything-goes, first vowel able to alternate with zero, initial consonant weak) are predicted to be also set to implement the two others.

This prediction is tested on a number of languages in Ségéral & Scheer (2008) and Scheer (2014a), showing its empirical substance. Whatever the ultimate result, though, the fact that a number of very specific phenomena occur at the left edge of words shows that the effect of this position is anything but arbitrary: it is not the case that, say, in some language word-initial consonants are especially strong, while in others they are especially weak; or that some languages restrict initial clusters to TR, while others allow only for RT. Rather, languages may or may not grant a specific status to the word-initial site. If they do not, word-initial consonants behave just as if they were word-internal. If on the other hand something peculiar happens at left edges of words, it is always the same phenomena that are observed. The following section explains why this invalidates diacritic carriers of morpho-syntactic information (such as # or ω).

11.1.3.3 Diacritics do not qualify

Since the 19th century, reference to morpho-syntactic information was always made by a diacritic, whose identity was determined by the basic phonological units of the time: juncture phonemes in structuralism when phonology was a string of phonemes, # which was held to be a [−segment] segment in SPE when the basic phonological units were feature matrices, and finally autosegmental domains such as the Prosodic Word ω or the Prosodic Phrase φ since the early 1980s when all areas of phonology were autosegmentalized.

By being a non-diacritic, the initial CV breaks with this tradition: syllabic constituents, an Onset and a nucleus, are not arbitrarily chosen and interchangeable symbols whose function reduces to the representation of morpho-syntactic information. If a # or an ω is replaced by an & and phonological processes then are said to occur in the vicinity of a banana, rather than of a # or an ω, the workings of the interface as we know it will remain unchanged. By contrast, a CV unit cannot be replaced by an & because it has a phonological identity which is independent from its eventual function of carrying morpho-syntactic information: an Onset is an Onset, not a banana; a nucleus is a nucleus, not a banana.

A number of arguments can be made against diacritic carriers of morpho-syntactic information (see Pak 2008: 60ff.; Samuels 2009: 284ff.; Scheer 2008b, 2012b: §93; 2014a: 316ff.). The most obvious of them is certainly the fact that diacritics are intrinsically unable to make predictions. Phonology does not react on the simple presence of a # or an ω – such items can only bear on phonology if the analyst has devised some instruction in phonological computation (a rule or a constraint) that is sensitive to them. Therefore hash marks, omegas and the like are passive 'sleepers': they merely sit in phonological representations without producing any effect by themselves.[8] For example, will a hash mark or a Prosodic Word favor or disfavor consonant clusters in their vicinity? There is no answer to this question because they can trigger (or inhibit) any phonological process and its reverse. By contrast, if phonologically meaningful items carry morpho-syntactic information, phonological computation will react on their bare presence. This is what Scheer (2012b: §154) calls the *Direct Effect*. A number of examples have been discussed above: the presence of the initial CV adds an empty nucleus to the string, a fact that has consequences: restrictions on initial clusters, inability of the first nucleus of the word to alternate with zero, strength of the initial consonant. These effects could not be the reverse: they depend on the phonological identity of the item inserted. Also, they are *automatic* and do not depend on any specific instruction: inserting #s and ω's is for free (no consequences need to be feared), inserting an empty CV unit is not: there is no way for an empty CV unit to land in the phonological string without being interpreted.

In conclusion, then, diacritics are disqualified by the non-arbitrary nature of the effects produced by morpho-syntactic information. Since the bare presence of diacritics does not have any effect and makes no prediction, representing boundary information by diacritics is claiming that anything and its reverse can be triggered by them. We know for sure that this falls foul of the empirical record.

Finally, another important difference between the insertion of syllabic space and regular units of the Prosodic Hierarchy is the linear character of the former. Syllabic space inserted (as much as structuralist juncture phonemes and SPE-type hash marks) becomes itself a piece of the linear string: it has a left and a right neighbor. An ω does not: units of the Prosodic Hierarchy define (autosegmental) domains, i.e. delineate a piece of the linear string to which phonological processes then make reference (domains of phonological computation). Modern phase theory (on which see section 11.1.4) does the same labour derivationally: it feeds phonological computation piecemeal with chunks that correspond to spell-out domains. It may therefore be asked whether grammatical theory can afford accommodating two distinct chunk-defining devices that do the same thing, one derivationally, the other representationally. Modular PIC, to be introduced in section 11.1.4.1, says no: there is only one chunk-defining device in grammar: spell-out (which defines computational domains in phonology).

11.1.3.4 Carriers of morpho-syntactic information reduce to syllabic space

Let us now zoom out to the global picture in order to see which units qualify for carrying morpho-syntactic information in phonology, and what their properties are. We have seen that diacritics do not qualify – hence the Prosodic Hierarchy has to go.[9] On the other hand, melodic primes, i.e. everything that is located below the skeleton (binary features, elements etc.), do not carry morpho-syntactic information either. While this fact does not appear to follow from anything, it is a paramount and consensual empirical generalization: the literature has not recorded any phenomena where melody would carry boundary information, nor has any theory devised melodic items as carriers of morpho-syntactic information: only items at and above the skeleton play this role. In GP, Bendjaballah (2012: 6) and Bendjaballah & Haiden (2013) have made this observation a central piece of their work. The observation itself is discussed in Scheer (2011: §660, 2012b: §124).

By elimination of those representational units that do not qualify for carrying boundary information, one concludes that *the output of translation of morpho-syntactic information* *reduces to syllabic space*. Note that this statement (as much as the list of items that do not qualify) is entirely theory-neutral. Depending on the theory favored, syllabic space may mean skeletal slots, moras, onsets, rhymes etc. In Strict CV, the minimal syllabic unit is CV.

The elimination of diacritics from the interface and the restriction of carriers to syllabic space are central pieces of Direct Interface theory (Scheer 2012b), which is called direct because there are no diacritic categories (#s, ω's etc.) mediating between morpho-syntactic information and phonological objects.

11.1.3.5 The initial and other CVs

Let us now look at some consequences of this setup. If the only object that is ever inserted into the phonological string in order to carry boundary information is syllabic space, it should be able to be the exponent of all kinds of morpho-syntactic values, not just of the beginning of the word.

This is indeed what the GP literature has found. Following up on Guerssel & Lowenstamm (1990), a line of research aims to identify the internal structure of templates (see Lowenstamm 2003). The idea is that templates of the Semitic kind are not just an amorphous set of consonantal and vocalic positions; rather, they have an internal structure. That is, morphological operations take place only on designated portions of the template.

Work along these lines includes Bendjaballah (1999); Bendjaballah & Haiden (2003); Lahrouchi (2001); Ségéral (2000); and Arbaoui (2010). Examples of boundary information that is found to be carried by CV units are the negative in Kabyle Berber (Bendjaballah 2001), a verbal marker in Chleuh Berber (Lahrouchi 2001) and little v, AspP and AgrP in Arbaoui's (2010) decomposition of Guerssel & Lowenstamm's (1990) Classical Arabic verbal template. Special mention needs to be made of the fact that work by Bendjaballah & Haiden (2003) has wandered outside of the Semitic or Afro-Asiatic family, showing that (much like Prosodic Morphology in the 1980s and 1990s) languages with predominantly concatenative morphology such as German may also possess portions of the syllabic makeup – CV units in their analysis – which are the exponents of specific morpho-syntactic information (such as tense in the ablauting system).

Cases of CV units other than the initial CV which carry morpho-syntactic information and are not related to (Afro-Asiatic) templates include the following. Charette (2003) and Luo (2013) hold that the right edge of words is marked by a CV unit in Turkish and Chinese, respectively. Pagliano (2003) argues that the exponent of a suffix class in French is a CV unit, which produces intrusive *t* as in *numéro-t-er* 'to number' and *bleu-t-er* 'to make something blue', as opposed to *bleu-âtre* 'bluish (pej.)': the infinitive *-er* comes with a CV unit, while the pejorative *-âtre* does not. In Italian, Passino (2008) argues that the non-nativeness of roots is marked by a CV unit: consonant-final roots, which do not occur in native vocabulary, geminate the root-final consonant in derivation (e.g. *tag* appears as *tagg-are* 'to mark'). Passino (2011) also analyzes a classical topic of Italian phonology, s-voicing, by recurring to the spell-out of an empty CV unit: in Northern varieties, intervocalic s is voiced before suffixes (/cas-ina/ → kaziina 'house, dim.') and before prefix boundaries (/dis-abile/ → dizaabile 'disabled'), but not after prefix boundaries (/a-simmetrico/ → asimmetrico 'asymmetric'). On her account, voicing is blocked when *s* is able to (covertly) geminate on an empty CV unit spelt out by the prefix boundary. Further, according to Passino (2014a: 20f.), so-called a-insertion in Italo-Romance dialects of the upper South amount to the marking of specific morpho-syntactic contexts by an empty CV unit (compare the final vowel of the word *kænə* in *lu kænə narə* 'the black dog' with *li kæna mi* 'my dogs').

11.1.4 Phase theory and Distributed Morphology

GP devices such as Element Theory on the one hand and results from phonologically oriented work on the interface regarding the initial CV on the other have been carried into the discussion of morphological and syntactic theory. Section 11.1.4.1 reports on which way the initial CV interacts with phase theory, while section 11.1.4.2 shows how work in Distributed Morphology (DM) of Jean Lowenstamm and his students uses Element Theory in the decomposition of morphological exponence.

11.1.4.1 What the initial CV is initial of

The study of external sandhi has shown that there are TR-only languages where the initial CV must be absent when cross-word phonology is computed. Belarusian (Scheer 2009b, 2012a: §285) and Corsican (Scheer 2009a, 2012a: §270) are cases in point. In the latter language, intervocalic stops lenite in external sandhi, i.e. when they occur in word-initial position before a vowel and the preceding word is vowel-final: voiceless stops voice (*um pánɛ – u bánɛ* ‘a/the bread’), while voiced stops spirantize (*un dntɛ – dui ðnti* ‘a tooth, two teeth’). The data mentioned also show that no lenition is observed when the preceding word ends in a consonant. This is unsurprising since the post-consonantal position is strong. In sum, the phonology across words behaves exactly as if it applied within words, i.e. as if the word boundary were not there. The presence of the initial CV would obliterate the alternations shown: word-initial consonants would always occur after an empty nucleus and hence be strong (in terms of the Coda Mirror) no matter what. Interestingly, though, consonants are also strong when occurring in a position where nothing can precede, i.e. utterance-initially and when quoted in isolation.

There are thus (at least) two patterns on record: in some languages, the initial CV is word-initial (no external sandhi), while in others it is utterance-initial. What the initial CV is initial to is thus variable: sometimes the word, at other times larger chunks (that exclude the word). That is, what languages mark with extrasyllabic space is initiality, i.e. the beginning of a given unit, whereby this unit may correspond to variable morpho-syntactic chunks.

If phase theory as currently entertained in syntax (Chomsky 2000 and following) is taken seriously, i.e. if it is believed to be the mechanism that organizes spell-out between morpho-syntax and phonology, morpho-syntactic structure can only impact phonology if it is spelt out. If on top of that we know that what is marked is initiality, the conclusion is that *what the initial CV is initial of is phases* (Scheer 2009a, 2012a: §307). Based on complementizer doubling and a-insertion after complementizers in Abruzzese (Italo-Romance), D’Alessandro & Scheer (2013) argue that more precisely what may be marked with a CV unit is the left edge of either the spell-out domain of a phase head (i.e. the complement of X°) or the phase head itself (X°).

This perspective prompts an issue for phase theory itself: phonological diagnostics for phasehood do not necessarily match syntactic diagnostics. In fact they rarely do, but they should if it were true that, as Chomsky (2000 and following) holds, phases are an active memory saving mechanism, and phonological computation needs to save active memory as much as syntactic computation. If it is true that in Corsican the CP is marked with an initial CV (because consonants are strong utterance-initially) but no smaller unit dominated by CP is (because, recall, that words cannot be preceded by an initial CV), it would be outlandish to conclude that Corsican as a whole has only one phase head. This phonological diagnostic does no justice to the syntactic workings: there is successive cyclic movement in Corsican. This and other standard syntactic diagnostics suggest that vP and maybe other functional categories are phase heads as well. They do not leave any trace in the phonology, though.

English offers established evidence that illustrates this mutual independence of phonological and syntactic footprints of phases. In American varieties, *t*-flapping is reported to operate across *all* word boundaries regardless of the syntactic relationship between the words (provided the *t* is word-final and intervocalic; see e.g. Nespor & Vogel 1986: 46f., 224ff.). Jensen (2000: 208) specifically mentions a case where flapping applies across a vP boundary: *a very dangerous wild ca[ɾ] escaped from the zoo*. On the other hand, there is firmly established syntactic evidence for the vP being a phase head especially regarding successive cyclic movement (e.g. Uriagereka 2011: 256). The vP in English is thus a case where a phase head leaves a footprint in syntax, but not in phonology. The reverse is also found: the assignment of word stress in English is strictly bound by the word, but the word is not a relevant unit in syntax, certainly nothing that would be described as a phase head for syntactic reasons.

D'Alessandro & Scheer (2015) show that all four logically possible configurations are found: a given phase head may leave a footprint both in syntax and in phonology, in neither module, or in one but not in the other. The result is *Modular PIC*: unlike in current phase theory, spell-out and the PIC are dissociated. There is a phase skeleton that defines phasehood for each language, i.e. at which points in the derivation spell-out occurs. An individual decision is then made for each access point whether a footprint will be left in syntax, and whether a trace of spell-out will be visible in phonology. In the latter case, footprints can be of two kinds, representational or derivational, and again all combinations are possible: spell-out domains, i.e. linear(ized) strings that reach phonology as an input, may or may not be associated with a PIC, and they may or may not be marked by an initial CV. Recall Kaye's example from section 11.1.2.1: in terms of Modular PIC, both nodes projected by class 1 and class 2 affixes trigger spell-out, but only the latter is associated to a PIC, i.e. freezes its sister. The conduit that organizes communication between morpho-syntax and phonology, spell-out, is thus the unifying spine of both representational and derivational management of the interface: the PIC and the initial CV may be associated to it.

In sum, Modular PIC is an attempt to bring phonological evidence to bear in order to develop phase theory into a general theory of the interface. The motor is previously unexploited phonological evidence: the pervasive mismatch of phonological and syntactic diagnostics for phasehood enforces more variable workings. Conversely, if phase theory is taken to be correct, a deeply rooted phonological mantra that transcends individual theories must be wrong: since Kiparsky (1982), postlexical phonology (i.e. phonology that applies across words) is supposed to be non-cyclic. According to phase theory, however, spell-out sends all kinds of strings that are larger than word size to phonological interpretation. It is implausible that a computational system, phonology, be entirely insensitive to its input conditions, i.e. never acknowledges the fact that packages arrive piecemeal.

11.1.4.2 Distributed Morphology and phonological exponence

Distributed Morphology (DM) does not feature much work related to phonology, and the items that include a phonological analysis often use minimal SPE-type vocabulary (e.g. Marvin 2002). Work by Lowenstamm (2008, 2011, among others) and his students (Rucart 2006; Arbaoui 2010; Lampitelli 2011; Faust 2013) does DM with stronger assumptions on the phonological side, and these are set in GP. Namely melodic representation in form of elements is concerned, since issues relevant to the organization of morphological structure involve exponence. At the heart of this strand is morphemic, rather than boundary, information. Lowenstamm (2011), for example, provides phonological arguments for the analysis of person, number and gender in Moroccan Arabic.

In order to illustrate element-based decomposition of morphemes, let us look at Lampitelli's (2011, 2013) analysis of Bosnian case markers (see also Passino 2014b; Lampitelli 2010 on markers of nominal inflection in Italian). Bosnian case markers are made of a vowel, which may be followed by a consonant and another vowel. Case markers express gender (roots belong to different declension groups according to their gender), number (singular and plural case markers are distinct) and case (six: nominative, accusative, genitive, dative, locative and instrumental). As is typical for Indo-European languages, all three values are expressed by one single morpheme, a vowel in this case, which at first glance appears to be indivisible: there is no piece of, say, the plural *-i* of masculine nominatives that corresponds to gender, number or case.

Lampitelli argues that this may be a wrong impression, though. Vowels decompose into elements, and these according to his analysis are the exponents of the three values to be realized. Case marking vowels, then, are compositional: they simply combine the elemental exponents of number, gender and case. The entire system is too intricate to be presented here, but let us look at a few prototypical cases. The unmarked values, which have a zero exponent, are (unsurprisingly) nominative, masculine and singular. Hence nouns with these values lack any case ending: *učenik* 'pupil'. The exponent of plural is an |I|, which produces the nominative plural *učenic-i*. The exponence of gender is as follows: masculine is zero, as was mentioned, feminine is |A| and neuter realizes |U|. Recall that gender produces three distinct declension classes. A feminine in nominative singular thus realizes two zeros (nominative, singular) and an |A|, which produces *kuć-a* 'house'. In its plural form, the plural marker |I| is added, to yield *kuć-e* where *-e* combines |A| and |I|. Neuter nouns display *-o* in nominative singular where *-u* is expected: *sel-o* 'village'. Since *-o* combines |A| and |U|, there is a supernumerary |A| whose origin needs to be accounted for. Lampitelli observes that there is syncretism between nominative and accusative in neuter nouns: *sel-o* is both nominative and accusative sg. Now accusative is marked by |A|, as shown by (animate) masculines: *učenik-a* (acc. sg.). Hence, Lampitelli argues, the case marker that is realized in the nominative of neuter singulars is in fact the accusative marker |A|, syncretically extended to the nominative.

While exponence in terms of specific elements and compositionality thereof produces correct results for a number of cells, the mechanism does not cover the entire paradigm. For those cells that resist, Lampitelli resorts to allomorphy rules such as Element → zero/__$A_{[GEN]}$, which erases all elements in presence of the genitive marker |A|. This accounts for the genitive plural of all three genders, which is uniformly *-a* (masc. *učenik-a*, fem. *kuć-a*, neut. *sel-a*): the rule erases the plural |I| in all three genders, as well as the neuter |U| in neuters. In genitive singular forms, nothing needs to be erased in masculines, which realize two zeros (gender and number) and the genitive |A| (*učenik-a*). The genitive singular of neuters, however, realizes the neuter |U| as well, and the expected output is |A|+|U|=[o]. Here again the allomorphy rule eliminates the neuter |U|, producing the attested *sel-a*.

More allomorphy rules are needed to account for the entire paradigm and additional (shallowly populated) declension groups. But the direction should be clear: following the general atomizing (i.e. anti-lexicalist) orientation of DM, a phonologically informed theory of exponence is able to shift labour from vocabulary insertion to the proper workings of phonology (and allomorphy), thereby achieving a one-to-one exponence, rather than the regular many-to-one exponence. That is, on the regular DM account, indecomposable case markers compete for realizing a portion of the tree that defines gender, number and case (many morphological features mapped onto one single exponent). The phonologically informed alternative assures that each morphological feature (or feature value) has its own phonological exponent. In sum, then, what is usually taken to be synthetic morphology may under such an account in fact be just as analytic as what agglutinating languages display overtly.

11.2 Interface with phonetics

11.2.1 Phonetic interpretation in Standard GP

The general and long-standing assumption concerning the interaction between phonology and phonetics in Standard GP is that phonological representations are directly mapped to phonetics. This rather enigmatic statement has gained new flesh recently, though the general outlook has remained. What has changed is that phonetic interpretation and inter-modular communication between phonology and phonetics has received more attention in recent publications in which the traditional phonetic interpretation of GP has been placed in a broader cognitive science environment. The recent developments are also consistent with evolutions outside GP (e.g. Hale & Reiss 2000, 2008; Hamann & Boersma 2009; Bermúdez-Otero & Börjars 2006), which build on phonology–phonetics mismatches and conclude that the relationship between the two domains is arbitrary. What has not changed within GP is that there are no computational steps within the phonological module towards a more concrete phonetic form. Rather, phonetic interpretation, or spell-out, is post-phonological and is done through lexical access (Scheer 2014b: 255).

How is it possible that phonological representations are interpreted phonetically without the mediation of computation that brings phonological representations to the level of systematic phonetic representation, which we know from SPE? The answer lies in the basic tenets of Element Theory (Kaye et al. 1985; Harris 1990, 1994, 1996; Harris & Lindsey 1995), which is part and parcel of Standard GP, and which is, with minor modifications, continued in recent incarnations of GP (cf. Chapter 9). Phonological elements and their combinations enjoy autonomous interpretability, that is, they are pronounceable without the need of further specification.[10] This means that phonological representations are fully interpretable regardless of the stage of phonological derivation we are in. In this sense, phonological derivation is not constructing representations that are any closer to phonetic representations. Therefore, there is in fact no need for a systematic phonetic representation (Harris & Lindsey 1995; Harris 2006, 2009). Harris (1994: 95) eloquently argues that the conception that phonological derivation turns more abstract phonological representations into concrete phonetic representations ‘assembling phonetic forms for production or reception’ places phonology outside the domain of generative grammar because then phonological knowledge would not be independent of performance. He argues that a truly generative role of phonology would be to turn phonological representations of some form into other well-formed phonological representations. More arguments along these lines can be found in Kaye (2005). But how do phonological categories relate to phonetic form?

Drawing on the Jakobsonian insight that grammar should be neutral between the speaker and the listener, Harris & Lindsey (1995) and Harris (1996) claim that phonological categories (elements) are first mapped onto acoustic signal, while perception and articulations are parasitic on this mapping. Harris (1996: 314) provides rough definitions of the universal set of phonological elements as gross acoustic patterns, which are idealized acoustic signatures. Similar views are expressed in Kaye (2005: 285), who maintains that phonological grounding is based acoustically rather than articulatorily, a view which is now also prevailing among phoneticians (e.g. Hamann 2011; Kingston 2007). Within GP, this point about the acoustic basis of phonological categories has recently been strengthened in Backley (2011) and Backley & Nasukawa (2009).

The above discussion may suggest that the universal set of elements have universally assigned phonetic (albeit only acoustic) substance. This is indeed the present-day 'official' position within Element Theory: universal association between phonological items (elements) and acoustic values, that is, *dip*, *rump* and *mass* for |I, U, A| respectively (Harris 1996). On the one hand, this facilitates talking about autonomous interpretation of representations. On the other hand, however, it takes us into a world of one-to-one universal relationship between phonetic (cues) and phonological (elements) categories. Below we will look at a substantial shift away from this position in recent studies which favor a view that the relationship between phonological and phonetic categories is in fact arbitrary, and established in language acquisition.[11] One of the consequences of this view is that the set of categories (elements) cannot be universal and innate. What humans have at birth is the ability to categorize physical input from the sensory system into cognitive units.

Despite the fact that not much has been said directly about the nature of phonetic interpretation within the GP tradition, individual proposals concerning concrete analyses provide a rather clear picture, which is expressed most emphatically in Kaye's (2005) *Phonological Epistemological Principle*, and which says that the only source of phonological knowledge is phonological behavior. From this it follows that the phonological representation cannot be successfully read off from the phonetic form, even if only acoustic cues are taken into account. Static phonetic properties may and do provide for initial hypotheses, which however can only be refuted or confirmed by phonological processing, which is the final judge. From the phonological practice, it also transpires that the rigid relation between phonological representation and phonetic exponence, whether acoustic or otherwise, needs to be relaxed as well.

For example, since the role of phonology, among other things, is to provide categorical phonological distinctions and the role of phonetic interpretation is to express these properties in concrete phonetic terms, one might be tempted to assume that two different phonological representations should never yield identical phonetic effects. One example that this view is wrong is provided in Gussmann's (2007: 56–61) analysis of *e*'s in Polish, which strictly follows Kaye's Epistemological Principle. In his analysis of palatalization, Gussmann observes two behavioral patterns in which *e*'s are involved in native vocabulary and concludes that despite identical pronunciation as [ɛ] there are in fact two different representations of this phonetic object. One of them is |I|-headed |A.I̲|. It palatalizes onsets, e.g. *rakiem* [racɛm] 'cancer, instr.', and is found word-initially after [j], e.g., *jest* [jɛst] 'is'. The second '*e*' is headless |A.I._| and does not palatalize onsets, e.g. *płotem* [pwɔtɛm] 'fence, instr.'. Thus, there are two representations of the front mid vowel which are realized in the same fashion, and it looks like phonetic interpretation has not fulfilled its obligation to express these phonological distinctions.

There is another way of looking at this problem, however. If phonetic interpretation, that is, the post-phonological spell-out, is a set of decisions independent of phonology proper, such mismatches are neither surprising nor problematic. Two disparate representations established on the basis of phonological behavior may receive the same interpretation, especially that they all involve a combination of the same two elements |I| and |A|. This, however, means that if there are any universal acoustic patterns associated with elements, they may be overridden by language-specific decisions. Note that from the perspective of language acquisition, when a child is confronted with data involving the same phonetic object behaving in two different ways, there is no choice but to give two different representations to that phonetic object.

Thus, the postulation of phonological categories depends in equal measure on the attempt to encode observable phonetic distinctions as phonological ones, as well as encoding phonetic non-distinctiveness with dual phonological behavior as two phonological categories. The guiding principle, however, is that one should observe phonological behavior. Some other examples of mismatches between phonetic categories and expected phonological categories will be mentioned below, and will be claimed to be due to arbitrary spell-out.

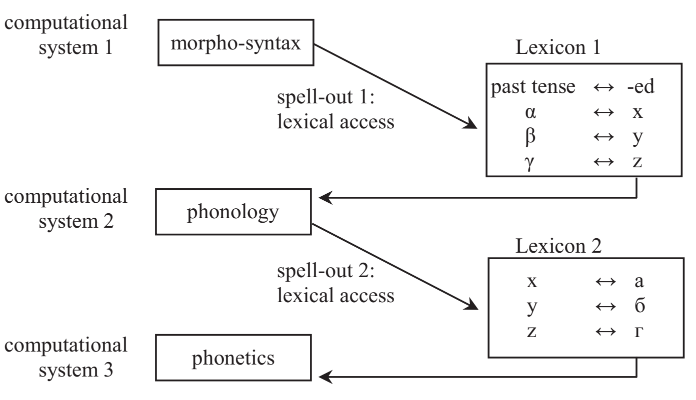
11.2.2 Two perspectives, one result: arbitrary spell-out

Below we show two distinct perspectives on the nature of the interaction between phonology and phonetics. One is global and theory-driven, while the other is data-driven but only possible under certain theoretical assumptions. They both converge on the same final conclusion: the relationship between phonological and phonetic categories is arbitrary.

11.2.2.1 Translation and arbitrariness

Working from a global modular architecture of grammar, Scheer (2014b) presents what the nature of the phonology–phonetics interface should be within GP (see also Chapter 9, section 9.1.2.1). One of the vocally articulated theoretical positions in this paper is that the spell-out must be viewed as arbitrary. The global modular architecture of grammar involving phonology is reproduced below (Scheer 2014b: 256).

(4) Fragment of grammar involving phonology



As encoded in the scheme, it is assumed that phonetics is a separate computational module that uses its own symbols and has a domain-specific battery of operations.[12] With this assumption in hand, the relationship of phonology with the lower end of grammar, that is, phonetics, must be of the same nature as with the upper end, that is, morpho-syntax (Scheer 2011). The communication between modules can take place only as translation (spell-out) because the vocabulary of two different modules are not mutually understandable, a point to which we return below. The spell-out is done through lexical access. It is a list-type conversion, very much like a dictionary list which is not subject to manipulation by any computation. This, in turn, suggests arbitrariness of the spell-out relations because lexical properties, or effects of translation, are as unpredictable as anything in a dictionary.

Scheer (2014b) notes a potential discrepancy between the nature of 'spell-out 1' above and 'spell-out 2'. While everybody agrees that morpho-syntax uses a distinct language from phonology and that this translation must be arbitrary – note the translation 'past tense' ↔ '-ed' in English: there is no reason why the exponent is '-ed' rather than, say, '-a' or '-t' – the distinction between the vocabulary used by phonology and phonetics, at least in terms of features, seems to be less obvious, and in fact more difficult to conceive of. For example, the phonological feature [labial] is sure to be interpreted as labial articulation with a corresponding specific acoustic signature. Thus, the domain of phonology–phonetics interaction is slightly more difficult to approach because of the similarity between the two modules. What is needed is a clear view with solid diagnostic criteria describing what the two modules are and what they can do.

One common misconception about the phonetic interpretation of phonological structure is that a production-oriented perspective seems to be implicitly assumed. This is inherent not only in the question: how are phonological representations interpreted phonetically, but also in the very terms 'spell-out', 'translation', or 'interpretation'. All these assume a directionality. This view enforces the use of similar if not identical vocabulary to talk about the two modules in question. Recall the association of elements to acoustic patterns in Harris (1996), mentioned above. This perspective forces us to say that phonological labiality translates into phonetic labiality or roundness. To see that this perspective is misleading it is sufficient to observe that such parallels are missing in the upper end of the grammar and one should also ask the question *why?* The past tense exponent '-ed' in English and '-ł-' in Polish are what they are for one single reason (excluding the historical development): these exponents have been lexicalized in acquisition. The spell-out connection has been established in acquisition, and is simply accessed each time a particular morpho-syntactic feature requires translation.[13]

Taking this into the phonology–phonetics domain now, the same mechanism can and should be expected and assumed. Since the phonological representation is established in language acquisition on the basis of phonetic input, including static patterns and distinctions between phonetic categories, as well as alternations, it is obvious that the feature [labial] or element |U| merely express the fact that this phonological category, whatever its real identity, has an established connection to labiality and its acoustic correlates. Thus, the similarity of the vocabulary used in phonology and phonetics may stem only from the fact that we do not know how to call the phonological category which is translated as labiality, so we use shorthand labels such as [labial], or |U|. Note that once we accept that phonetic interpretation, or spell-out is a case of access or activation of the relations already established in acquisition, the problem of autonomous interpretability of elements, and more broadly, of interpretation of truly phonological representation, vanishes. All we need to focus on more is the nature of the relations and criteria for deciding what is a truly phonological process, what is a phonetic phenomenon and what are the principles of interpretation. In other words, how the phonological representations are established. The story of Polish '*e*' is one example of this. To illustrate this point further we may use another example, that of Russian '*v*'.

11.2.2.2 Arbitrariness and the nature of mismatches

The behavior of Russian '*v*' has been discussed in the literature on many occasions (e.g. Andersen 1969; Hayes 1984; Mołczanow 2008). The problem with this segment is that it sounds like an obstruent and behaves like a sonorant in some contexts and like an obstruent in others. The typical line of analysis is that it is an underlying sonorant which is turned into an obstruent by derivation. The problematic nature of this speech sound, it seems, is due to the assumption criticized above, that there is a one-to-one correspondence between the phonetic cues and phonological representation. Under the view that phonological representation is primarily governed by phonological behavior (Kaye 2005), the Russian '*v*' is no more problematic than the Polish '*e*'.

The small percentage of cases where we observe a mismatch between phonological and phonetic categories is usually due to the procedural rather than static considerations. In other words, observable phonological processes, such as alternations, determine the actual phonological representation of seemingly straightforward phonetic objects. Surely there is nothing in principle that should preclude lexicalization of the input [v] as a phonological object involving the categories for labiality |U|, friction |h| and voicing |L|, to use a Standard GP application of Element Theory, unless some evidence, process or alternation tells us otherwise. The Russian '*v*' is a classic example of this.[14] It is best represented by |U| or |<u>U</u>| only. The friction as well as voicing are non-phonological. What is more, as argued in Cyran (2014b), obstruentization is not only unnecessary but also impossible as a synchronic phonological process. The need to call Russian '*v*' a voiced labio-dental fricative (an obstruent) stems only from the assumption that all the phonetically distinctive properties must find a reflection in the phonological representation. The error of this thinking lies in the fact that it ignores the Phonological Epistemological Principle of Kaye (2005).

Let us now return to the question posed in Scheer (2014b), namely, why should there be so much one-to-one correspondence between phonological and phonetic categories if the relationship is arbitrary (see also Hamann 2011, 2014). We are able to say that partly this one-to-one correspondence is a linguist's illusion. Firstly, if we reanalyze all familiar melodic phenomena using the Phonological Epistemological Principle and Element Theory, we might conclude that the ratio is not so overwhelmingly in favor of one-to-one relationship. And secondly, given the arbitrary spell-out the phonological categories need not correspond to phonetic labels at all. In the extreme case the phonological elements could be numbers, shapes or colors. Current Element Theory has not gone that far. However, it seems to do enough to remove phonological categories from phonetic ones. Recall that Harris defines elements in terms of gross acoustic patterns. In this sense, we could say that phonological elements are idealizations of the acoustic cues they have been related to in acquisition.

Scheer (2014b) gives yet another explanation for the general one-to-one correspondence as well as the observed mismatches. In his view, this overall picture follows from diachronic development in which phonetic faithfulness is present only in the case of fresh lexicalizations of phonological processes, and it may wane by aging and rule telescoping.

11.2.2.3 Modular constraints on translation

In this section we focus on constraints on translation which follow from the modular organization of grammar shown in (4) above (Scheer 2014b: 258–260). Thus the general constraints established on the basis of the interface between phonology and morpho-syntax will be enumerated and applied to the translation of phonological output into phonetic alphabet.

One of the properties is that spell-out through lexical access must take a form of *list-type conversion*. This means that phonological categories are lexically associated with particular phonetic categories and stored in the long-term memory. This does not preclude diachronic change, however. We should probably add a speculation that there is no size limitation as to how much phonological structure is associated to a given phonetic exponent. This point may look a little unconstrained and probably requires some more explicitness. However, there is ample evidence from existing studies within GP in which chunks larger than one segment are subject to a phonetic interpretation suggesting that we are dealing with a single segment or simplex representation. Scheer (2014b: 263) calls this phenomenon virtual length, which is a case in which phonological length that is typically interpreted in phonetics as duration may sometimes be spelt out differently. One example of this situation is the English angma. The velar nasal [ŋ] has been shown to be a phonological cluster /ng/ (e.g. Gussmann 1998). Numerous studies within GP also show that in some languages phonological length of vowels is not distinctive phonetically. For example, in Apulian dialects of Italian (Bucci 2013a, 2013b) phonological length of vowels translates as non-reduction, while short vowels are reduced to schwa. Finally, there are studies which demonstrate that consonantal length is interpreted as shortness of the preceding vowel (Caratini 2009; Cyran 1996; Ségéral & Scheer 2001).

Another property of inter-modular translation is its necessarily *non-computational* nature. The very fact that we are dealing with two distinct modules precludes a computational translation from one system to the other. In our view, contrary to the dominating outlook in the phonological tradition since SPE to this day, it is impossible to imagine how some computational system would turn morpho-syntactic features into phonological categories, or a phonological system into phonetic categories for that matter.

Arbitrariness is a consequence of inter-modular translation that we have already mentioned above. What needs to be added here is the perspective from which we need to look at it. The arbitrariness that we observe at the interface between morpho-syntax and phonology mentioned above is never a problem for acquisition or spell-out. The arbitrary relation between the phonological form and the meaning in German *Hund* and Polish *pies* for 'dog' is simply established in acquisition. Thus, we need to bear in mind that the actual synchronic state of the grammar at any level, including the phonology–phonetics interface, came into being from an opposite direction to production. The original decision in the construction/internalization of grammar was to connect exponents located in the input to particular features of grammar at higher levels. Thus, there is never a problem with externalization of a previously internalized grammar. Spell-out does not decide on anything; it is merely an activation of an existing, previously formed connection. In this sense, arbitrariness is also fully compatible with Kaye's Phonological Epistemological Principle. The connections need not be one-to-one.

A final property that follows from the general behavior of translation is *exceptionlessness of conversion*. Lexical relations are never mistaken or prone to error. Just like at the upper interface: there is 100% regularity of the match between past tense and '-ed'. Here we observe a convergence of this property as diagnostic of spell-out, that is, phonetic interpretation of phonological representation, as well as of truly phonological processes. Note that, like in Natural Generative Phonology (e.g. Hooper 1976), GP also claims that processes which have exceptions must not be viewed as phonological. The obvious question then is how to tell the difference between 100% regular phonological processes and 100% regular spell-out effects, which will also amount to alternations in some cases. The dilemma is not trivial and requires further study. We need to know how to distinguish the two types of phenomena, as well as to know for what reasons. It may turn out that this ambiguity is not unwelcome.

To illustrate the ambiguity described above we may briefly look at the alternation [v~w] in Belarusian (Scheer 2012b: 223–231).[15] Within a word, the distribution of the two allophones is clear and can be described in traditional terms as [v] occurring in the Onset, that is, pre-vocalically, and [w] in the 'coda' position, that is, pre-consonantally and word-finally, which in GP terms translates as: in front of an empty nucleus (cf. Chapter 10.2.3.3). Thus, in *korov-a* 'cow, nom.sg.' and *vad-a* 'water' we get [v], while in *korow-ka* 'cow, dim.nom.sg.' and *korow* 'cow, gen.pl.' the glide [w] is found. To a great extent, given the conflation of regular spell-out and regular phonology, the descriptive problem with this allophony resembles that of English aspiration, namely, two distinct analyses can be offered. If the alternation is phonological, some phonological computation must be assumed to derive it. One simplified analysis will be shown presently. If, on the other hand, the alternation is purely interpretational (a case of post-phonological spell-out), then no derivation of /v/→/w/, or /w/→/v/ should be assumed. Rather, we would be dealing with a spell-out of the same representation as two distinct phonetic objects in the relevant contexts, say, |U|↔[w]/_ø, and |U|↔[v]/_V. That such an interpretational shift must have occurred in most Slavic languages is argued for in e.g. Cyran (2014b), who additionally eschews |U|→|$\underline{U}$| as a possible phonological process of strengthening. On the other hand, loss of headedness under weak licensing |$\underline{U}$|→|U| is licit. Below we compare the two stories: phonological and interpretational.

In the phonological analysis, we follow Scheer's assumption that the underlying representation is /v/, that is, headed-|$\underline{U}$|. In (5a) below, we observe the derivation of the weak [w], which, to simplify things a little, may be said to occur under weak licensing of the following empty nucleus and takes the form of loss of headedness. The pre-vocalic context in (5b) provides sufficient licensing for the headedness of |$\underline{U}$|, therefore the result of computation is the same as the lexical representation. (5c), on the other hand, shows the same alternation viewed as a case of spell-out. This time, since there is no phonological computation, it does not matter if the underlying segment is |U| or |$\underline{U}$|. We go for the headless one. It should be noted that the point that is made here is entirely independent of the phonological theory used. Whether we use the Element Theory in GP, or a different model, the phonological/computational analysis will involve a syllable-based change of phonological identity of /v/ to /w/ in the coda (or in front of an empty nucleus in GP), while in the spell-out analysis, the same phonological object undergoes context-sensitive spell-out, not a change of identity. Below in (5) we provide phonological representations of whole words, but the focus is on the element |U|. Note that the role of context in the phonological analysis (5a,b) is instrumental in computation, but irrelevant for spell-out in that it takes a list-type translation: |$\underline{U}$| ↔ [v], |U| ↔ [w]. On the other hand, in the spell-out analysis (5c), the context plays a role in distinguishing the two phonetic outcomes of the same phonological object: |U| ↔ [w]/_ø, |U| ↔ [v]/elsewhere.

(5)	Phonological representation and computation				Spell-out	
a.	/koro$\underline{U}$ø/	→	/koroUø/	↔	[korow]	|U| ↔ [w]
	/koro$\underline{U}$øka/	→	/koroUøka/	↔	[korowka]	|U| ↔ [w]
b.	/koro$\underline{U}$a/	→	/koro$\underline{U}$a/	↔	[korova]	|$\underline{U}$| ↔ [v]
c.	/koroUø/			↔	[korow]	|U| ↔ [w]/_ø
	/koroUøka/			↔	[korowka]	|U| ↔ [w]/_ø
	/koroUa/			↔	[korova]	|U| ↔ [v]/elsewhere

There is no *a priori* reason why the computational analysis should be superior to the interpretational one. In fact, it is argued in Cyran (2014b) that (5c) must have been the initial stage of the strengthening shifts involving the Common Slavic **w*. It was interpretational at first, and then phonologized.[16] (5a,b) are cases of phonologization of that shift, turning it into a computational phenomenon and allowing for further shifts including a lexicalization of the sonorant-like [v], that is, |$\underline{U}$|, as the full-blown obstruent /v/, that is, |U.h.L| in some modern Slavic languages, e.g. Polish. For the shift from (5c) to (5a,b) to occur, a phonological condition must be fulfilled. Namely, the distribution of strong (headed) and weak (headless) objects must correlate with prosodically defined strong and weak positions, to ensure that the distribution of headedness is phonologically/computationally non-arbitrary.

To conclude, the ambiguity between 100% regular spell-out and 100% regular phonology is in fact a welcome situation because it describes the conditions for diachronic change to occur: it is a phonologization of shifts that originate in spell-out.

So far, we have looked at the global modularity-based argument for post-phonological spell-out, and have added some empirical flesh to it as well as shown some consequences. Below, we look at how very much the same conclusions have been reached from a data-based end, which, however, could not be possible if a particular strict version of Element Theory was not assumed.

11.2.2.4 Laryngeal relativism

Apart from the global perspective of modularity and cognitive science presented in Scheer (2014b), one may argue for the arbitrariness of the phonology-phonetics spell-out from a theory-specific point of view, coupled with empirical solutions (Cyran 2011, 2014a). What is rather impossible and futile, as will transpire below, is approaching the spell-out problem from a purely empirical, data-based or 'phonetic-facts-based' position.

Laryngeal phonology is one of the few areas in the field where similar representational views seem to be held across frameworks. The main one is privativity of laryngeal categories. In recent years, laryngeal realism (Iverson & Salmons 1995; Honeybone 2005; Harris 2009) seems to have gained wide acceptance. It assumes, for example, that languages with a two-way laryngeal contrast divide into two different systems with different categories involved in the privative marking. The so-called 'voicing' languages (e.g. Slavic and Romance) oppose fully voiced stops with voiceless unaspirated ones, while 'aspiration' languages (typically Germanic) contrast voiceless aspirated stops with the voiceless unaspirated or passively voiced ones. The three phonetic categories – that is, fully voiced, voiceless unaspirated and voiceless aspirated – form a well-known continuum along the VOT dimension (Lisker & Abramson 1964). The guiding principle in laryngeal realism is that phonological categories correspond to the members which constitute a VOT displacement from the neutral, that is, voiceless unaspirated. Thus, for example, Polish voiced obstruents must be defined by the presence of the element |L|, which corresponds to traditional [+voice]. We will symbolize such voiced obstruents as C^L. On the other hand, the voiceless unaspirated series are neutral (unmarked), that is, C^o. In contrast, English fortis series are C^H, where the element |H| roughly corresponds to [spread glottis], and are opposed to C^o. Within GP, laryngeal realism is widely recognized in Harris (1994, 2009); Honeybone (2005); and Gussmann (2007).

In the light of our discussion above, it is immediately obvious that laryngeal realism, although it constitutes a welcome advancement in the theory of subsegmental representation, is not exactly compatible with the expected arbitrariness, which is one of the offshoots of post-phonological spell-out stemming from an independent modular status of phonology and phonetics introduced above. The one-to-one relationship between the phonetic cue 'negative VOT' and the representation of that cue as element |L| is anything but arbitrary in nature. The assumption that, when met with a two-way contrast [b-p] in acquisition, the child will automatically assign a marked status to the voiced obstruent, rather than to the voiceless one, is also incompatible with Kaye's Phonological Epistemological Principle. What if the phonological behavior of the objects in question suggests a reversed marking? One such situation is discussed below.

Laryngeal realism seems to fail when confronted with a number of phenomena. One case in point is Polish, in which thus understood laryngeal realism allows us to understand only one of the two main dialect groups, and leaves us helpless with respect to the celebrated phenomenon of Cracow-Poznań (CP) sandhi voicing. In order to fully appreciate the arguments below, we need to be clear about a basic theoretical assumption of Element Theory which we wish to strictly adhere to, namely, privativity. This means that two-way contrasts are represented by the presence versus absence of one phonological category (element). And more importantly, the absence of contrast means that there is no marking. The latter situation refers to sonorants, that is, vowels and sonorant consonants, which are neutral from the point of view of laryngeal specification (V^{o}, S^{o}), and their voicing is spontaneous.

The main voicing phenomena in Polish, to simplify things a little, are final obstruent devoicing (FOD) and voicing assimilations (VA) between obstruents. Both processes involve phonological computation in the form of delaryngealization (element deletion) before an empty nucleus, and spreading. Phonetic interpretation also seems to play an important role in the phenomena. Given the assumption of laryngeal realism, the voiced obstruents in Polish have |L|, and the phenomena can be described in the following way. The alternation *waga/wag* [vaga ~ vak] ‘scale, nom.sg./gen.pl.’ with FOD involves loss of |L| before an empty nucleus: /vag^{L}a/ ~/vag^{L}ø/ → /vag^{o}ø/ ↔ [vak]. Recall that the laryngeally neutral obstruent must be spelt out as [k]. One of the regressive assimilations also involves a similar mechanism. Namely, in the alternation *kawa/kawka* [kava ~ kafka] ‘coffee, nom.sg./dim.’, the only phonological operation involved is also delaryngealization of /v^{L}/: /kav^{L}a/ ~ /kav^{L}øk^{o}a/ → /kav^{o}øk^{o}a/ ↔ [kafka]. Note that this assimilation is not due to spreading because the following obstruent is neutral as well and has nothing to spread. Thus, the phonological computation is limited to delinking of |L|, while the actual assimilation is merely interpretative, that is, a case of spell-out. Spreading is present in the alternation *prosić/prośba* [prɔɕit͡ɕ~ prɔʑba] ‘to ask/a request’ in that |L| spreads from /b^{L}/ leftwards: /prɔɕ^{o}it͡ɕø/ ~ /prɔɕ^{o}øb^{L}a/ → /prɔɕ^{L}øb^{L}a/ ↔ [prɔʑba].

The phenomena of FOD and VA, as well as the existence of the voiced/voiceless distinction, are uniform in all dialects of Polish. The differences come out in the so-called pre-sonorant CP sandhi voicing, which is not observed in Warsaw Polish (WP).

(6)		WP	CP	
a.	*grup otwartych* ‘open groups’	p-ɔ	b-ɔ	_V^{o}
	*gró*b *otwarty* ‘open grave’	p-ɔ	b-ɔ	
b.	*grup matek* ‘groups of mothers’	p-m	b-m	_S^{o}
	grób matek ‘grave of mothers’	p-m	b-m	
c.	*grup dorosłych* ‘groups of adults’	b-d	b-d	_C^{L}
	grób dorosłych ‘grave of adult’	b-d	b-d	
d.	*grup takich* ‘such groups’	p-t	p-t	_C^{o}
	grób takich ‘graves of such’	p-t	p-t	

We can immediately eliminate (6c,d) from our discussion as the results are the same in both dialects. Within laryngeal realism, (6c) is a case of |L|-spreading from the following C^{L}, while (6d) is due to |L|-deletion. Let us note that just as in all the other examples in (6), we are dealing with a neutralization of the laryngeal distinction in word-final context, in that the lexical distinction is lost.

(6a,b) appear to be fully predicted as far as WP is concerned. Given that |L| is neutralized in the word-final context, and the following sonorants do not possess a spreadable laryngeal element |L|, it is expected that the final obstruents will be uniformly voiceless. This, however, is not what happens in CP. Both types of obstruents, that is, lexically voiced and lexically voiceless, are voiced in front of vowels and sonorant consonants.

It is clear that we are not dealing with mere retention of |L| in these forms, because that would concern only the lexically voiced obstruents. The alternative solution, then, must be that voicing comes from the following sonorants. This is problematic for laryngeal realism because we suddenly have to admit that sonorants may be marked with |L|, but only in CP. Needless to say, this is not even an option within the strict version of Element Theory (Harris 1994).

An alternative solution was proposed in Cyran (2011, 2014a), which retains strict privativity and non-marking of sonorants and applies the Phonological Epistemological Principle in contravention of the principles of laryngeal realism. Under this new proposal, given the two-way contrast [b-p], the choice of which series is to be marked depends on the phonological behavior alone. This leads us to an inverted laryngeal marking in CP, one in which the full voicing is a spell-out of an unmarked object, while the voiceless object has |H| as in typical ‘aspiration’ languages, except that with no aspiration.

Let us first see how this reversed system handles the familiar processes of FOD and VA. It appears that the interpretative system must be quite different from that of WP, while the phonology, except for the inverted marking, remains identical. FOD in the alternation *waga/wag* no longer involves delaryngealization. It is simply a case of absence of passive voicing which is observed in /vag^{o}a/ ↔ [vaga] in phonetically non-voiced environment: /vag^{o}ø/ ↔ [vak]. Thus, the new situation here is that FOD is phonological in WP, but interpretational (spell-out), and thus in a sense phonetic, in CP. Assimilations are surprisingly non-problematic in this inverted system. The alternation *prosić/prośba* involves regular delaryngealization in front of an empty nucleus: /prɔɕ^{H}øb^{o}a/ → /prɔɕ^{o}øb^{o}a/ ↔ [prɔʑba], cf. /prɔɕ^{H}it͡ɕø/ ↔ [prɔɕit͡ɕ]. This is a case of phonetic/interpretational assimilation (passive voicing). Finally, the alternation *kawa/kawka* is a case of absence of passive voicing in front of a phonetically voiceless obstruent, which is marked to be so: /kav^{o}a/ ↔ [kava] ~ /kav^{o}øk^{H}a/ ↔ [kafka]. It is possible to talk about |H|-spreading here, but it is not even necessary. The neutral obstruent requires a following voiced context in order to be pronounced voiced in this dialect.

We saw in the sandhi data that the word-final context neutralizes the laryngeal distinction. This means that in WP the element |L| is lost, while in CP it is |H|. Both dialects end up with a neutral C^{o} in that context. However, these are systemically different animals. In WP, C^{o} may be voiced only if |L| is spread from the following word, as we see in (6c). In CP, on the other hand, all that is required now is a phonetically voiced segment in the following word and the neutral C^{o} should be spelt out as voiced through passive voicing, which it does in (6a–c). It is important to realize that this analysis does not require any rule of CP sandhi voicing. The phonetic interpretation of C^{o} in sandhi is exactly the same as word-internally, that is, $C^{o}V = C^{o}\#V$, $C^{o}S = C^{o}\#S$, $C^{o}C^{o}V = C^{o}\#C^{o}V$, and $C^{o}C^{H}V = C^{o}\#C^{H}V$.

Thus, it is possible to provide an analysis of CP sandhi voicing without compromising strict privativity and non-marking of sonorants. There is also no need for rule ordering (Rubach 1996). The phonological computation (delaryngealization and spreading) naturally precedes spell-out. All that we did was apply all principles of GP and Element Theory strictly, including the Phonological Epistemological Principle. But, as a consequence, we need to break with laryngeal realism. The phonological marking of laryngeal contrasts is not given directly in the signal. The signal provides information that we need a category to distinguish two series, as well as plenty of information concerning the behavior of these two series. It is the latter type of information that determines the type of marking, which happens to be reversed in the two dialects of Polish.

As with the two alternative analyses of [v~w] in Belarusian, discussed above, we observe similar consequences of the presence of post-phonological spell-out. One of them is the ambiguous nature of phonetic facts. Voicing can be phonological, with an active phonological category which will participate in phonological processing such as deletion and spreading, but it may also be a result of a particular systemic spell-out, in which case it is more phonetic in nature, and phonetics-dependent. FOD is phonological in WP and interpretational in CP. Assimilations can be due to spreading, but also due to interpretation. Thus, all these phenomena must be treated with caution, and representational conclusions must not be drawn on the basis of phonetic properties alone. Such is laryngeal relativism.[17]

11.2.3 Further issues and perspectives

This section is to some extent speculative. One of the main aims for future research within the program sketched above, in which phonology and phonetics are separate modules and communicate through arbitrary translation, is first to delineate the two linguistic modules and define their disparate characteristic behavior, as well as determine the principles of translation, if there are others than the ones discussed in section 11.1.2.2 above. This is not an easy task. One reason for this difficulty lies in the fact that the delineation should be radical if it is going to bring any results.[18] Otherwise, the boundary between phonology and phonetics will continue to be unclear. For example, the assumption of substance-free phonology, for which the presented model of Element Theory seems to be cut out, does not seem to loom on the horizon even though this should be the very first step in order to move on. GP in general is also best suited to explore such a path.

The phonological module as practiced in current versions of GP is much smaller than generally assumed in other models (cf. Chapter 9, section 1.3). It simply involves mostly syllabic representation with privative elements and very restricted computation limited to the arrangement of government and licensing and a small number of melodic operations such as decomposition (e.g. lenition, that is, delinking of elements) and composition which must involve spreading of a property from a local source, e.g. |L|-spreading in voicing assimilation, or resonance element spreading in vowel harmony. The only phonetic presence in this theory is the acoustic definition of phonological elements, and possibly, the very division of the skeleton into Cs and Vs.

The definition of phonetics as a computational module is not simple either. A number of universal principles which can be harnessed to explain linguistic sound systems by providing, for example, the rationale for particular phonetic categories used as spell-out targets in phonology–phonetics translation, may be claimed not to be phonetic, but rather belonging to more general cognitive strategies parallel to other non-linguistic ones. The quantal theory of Stevens (1972) or the dispersion theory (Liljencrants & Lindblom 1972; Schwartz et al. 2007) are interesting proposals within phonetic theory, but are they really talking about phonetics-specific properties of the human brain? Inference in perceptual studies might be viewed as a good candidate to pass for computation in the phonetic module (Reiss 2007), but is it really only phonetic? The same principles hold in visual perception. The 'phonetic' nature of voicing is in fact physics (aerodynamics). Thus, our views on phonetics as a linguistic module must also crystalize. It is possible that the grammar–non-grammar boundary runs between phonology and phonetics. This does not undermine the scheme in (4) showing inter-modular communication. It just tells us that the translation between phonology and phonetics is still more complex than we envisage today.

11.3 Further reading

Kaye, Jonathan 1995. Derivations and interfaces. *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 289–332. London & New York: Longman. (Also in *SOAS Working Papers in Linguistics and Phonetics* 3, 1993, 90–126.)
Introduces the concepts of domain structure (akin to cycles) and analytic vs. non-analytic morphology (akin to class 1 vs. class 2 affixes in English). Morphological boundaries are either visible (analytic) or invisible (non-analytic) to phonological computation. The article introduces what will later be known as Phase Impenetrability (robustness) and the Phase Edge (the sister of a phase-relevant node is spelt out, rather than the node itself).

Scheer, Tobias 2014. The initial CV: Herald of a non-diacritic interface theory. *The Form of Structure, the Structure of Form: Essays in Honor of Jean Lowenstamm*, edited by Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi & Nicola Lampitelli, 315–330. Amsterdam: Benjamins.
Lowenstamm (1999) has introduced the idea that morpho-syntactic information may incarnate into a truly phonological object that exists anyway (a CV unit), rather than into a diacritic (#, ω). This article explains from hindsight how the initial CV has paved the way of a non-diacritic interface theory.

Scheer, Tobias 2012. *Direct Interface and One-Channel Translation: A Non-Diacritic Theory of the Morphosyntax-Phonology Interface: Vol.2 of A Lateral Theory of Phonology*. Berlin: de Gruyter.
The book introduces and motivates a non-diacritic theory of the interface, i.e. where phonological carriers of morpho-syntactic information are true phonological objects that are also used beyond interface issues. This theory is theory-neutral, i.e. may be implemented into any individual phonological theory. Its incarnation using the representational vocabulary of Strict CV is described.

Lampitelli, Nicola 2013. The basic elements of inflection: Morphophonology of Bosnian nouns. *Formal Approaches to Slavic Linguistics 20: The Second MIT Meeting 2011*, edited by Alexander Podobryaev, 154–170. Ann Arbor: Michigan Slavic Publications.
Exposes and illustrates the idea that morphological units may spell out as subsegmental phonological items (elements), rather than as full vowels, consonants or combinations thereof. Hence in Bosnian, the nominative plural of the feminine declension is [-ɛ] and decomposes into A (feminine), I (plural) and zero (nominative).

Harris, John 1996. Phonological output is redundancy-free and fully interpretable. *Current Trends in Phonology. Model and Methods*, edited by Jacques Durand & Bernard Laks, 305–332. Salford, Manchester: ESRI.
Presents a clear picture of phonetic interpretation in GP and argues against the view that phonological derivation produces representations which are closer to the systematic phonetic level.

Harris, John 2009. Why final obstruent devoicing is weakening. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 9–45. Berlin and New York: Mouton de Gruyter.
The paper is a good example of a dominant philosophy concerning the relationship between phonology and phonetics in GP, placing emphasis on extracting phonological information from the signal.

Scheer, Tobias 2014. Spell-out, post-phonological. *Crossing Phonetics-Phonology Lines*, edited by Eugeniusz Cyran & Jolanta Szpyra-Kozłowska, 255–275. Newcastle upon Tyne: Cambridge Scholars Publishing.
This programmatic paper places the phonology–phonetics interaction in a broad context of all intermodular communication in language, arguing, for example, for arbitrariness of spell-out.

Cyran, Eugeniusz 2014. *Between Phonology and Phonetics: Polish Voicing*. Berlin: Mouton de Gruyter.
This is a book-length study of voicing phenomena in Polish from the perspective of the phonology–phonetics interaction. It argues for an arbitrary relation between these two domains.

Notes

1 Scheer (2011: §26, 2016b) provides discussion of the general environment, including approaches that lie beyond the modular frame.
2 See Nasukawa (2011, 2016) for an approach called Precedence-Free Phonology where linearity does not pre-exist phonological computation (upon production) but follows from dependency relations among the units of phonological hierarchical structure.
3 Though, following SPE, with an implicit recognition of a specific treatment of the word level (see Scheer 2011: §338).
4 See Scheer (2011: §271) for more discussion regarding the combination of concat and φ, Gussmann (2002: 45ff.) for a general introduction to domain structure.
5 A more detailed review of no-look-back devices that the literature has accumulated since 1973 is available in Scheer (2011: §287).
6 Scheer (2014a) provides an overview of the offspring of this idea, also including a more detailed discussion of the material covered below.
7 Kula & Marten (2009) discuss the strength of word-initial consonants in languages that lack word-initial clusters.
8 The history of diacritics in Prosodic Phonology is documented in Scheer (2011: §365): in the early 1980s linear diacritics (SPE-type #'s) were replaced by autosegmental diacritics (ω etc.).
9 To be precise: the higher levels of the Prosodic Hierarchy from the Prosodic Word ω on have to go. Feet, syllables and moras are different because they are bottom-up constructions, i.e. the projection of basic units. By contrast, the Prosodic Word etc. is the projection of nothing (see Scheer 2012b: §138).
10 Apart from elements and their combinations, phonological structure can also be directly interpreted, except for empty positions. For example, association of a resonance element to two skeletal positions (long vowel) may produce a tenser variety than the same element linked to just one position (cf. English [ɪ] vs. [iː]). Coda–onset governing relations of the Standard GP type have also been assumed to be interpreted as 'stopness' leading to the elimination of the occlusion element |ʔ| (Jensen 1994). In more recent proposals under the banner of GP 2.0, the former element |A| is expressed phonologically as a subsegmental tree structure akin to syntactic ones (Pöchtrager & Kaye 2013). The latter proposal is very much in the spirit of standard Element Theory in that a universal one-to-one connection between representation and phonetic interpretation is assumed, an idea that we attempt to dismantle in this chapter.
11 Although the emergent nature of elements within GP is not an accepted view, it will be shown that arbitrariness of spell-out leads directly to this conclusion. This does not mean that elements cannot be defined in terms of universal acoustic patterns. All that it means is that the relation is not innate (e.g. Mielke 2008). The supposed universality is due to phonetic and functional factors such as the characteristics of the vocal tract and perception.
12 The claim that phonology and phonetics are separate modules will not be argued for here. It follows from the phonological model assumed here, that is, GP, which is sharply distinguished from phonetics. It is also an open question what counts as phonetics and what phonetic computation might look like and if it is indeed necessary to conceive of such a module. The absence of computation in phonetics would appear to weaken the global picture in which each module is characterized, among other things, also by possessing its own domain-specific battery of computational operations.
13 For this reason, the symbol '↔' to refer to spell-out, which is used in Scheer (2014b), touches the heart of the matter. The translation relation is bi- if not non-directional. It is static in a fully developed grammar, and facilitates parsing in perception as well as articulation in production. Note that at the acquisition stage, it is in fact the reverse direction from the one assumed in the production-oriented perspective: the phonological representation is established on the basis of phonetic input (phonology ← phonetics).
14 For a general Slavic perspective on the labial glide strengthening within GP, see e.g. Cyran & Nilsson (1998) and Cyran (2014b).
15 In fact the alternations involve [u~w~v] and are much more complicated than presented here. For example, the discussion in Scheer (2012b) aims to capture the phonological behavior of these allophones with respect to word boundaries. We will limit ourselves to the word-internal situation. This, however, has no consequence on the argument in question.
16 The reason why strengthening [w]>[v] could not be phonological at the initial stage follows from a strict application of Element Theory, which limits potential phonological processes to

decomposition (element loss) and composition (addition of elements and properties like headedness) as a result of spreading. Thus, obstruentization as a synchronic process is also ruled out because there is no local source of spreading of headedness. Obstruentization may only occur as a phonologization of a spell-out pattern.

17 See van der Hulst (2014) for an analysis of Dutch in the spirit of laryngeal relativism. His conclusion is not only that Dutch (a 'voicing' language by laryngeal realism standards) is an |H|-system like CP, but he also proposes that all two-way systems are |H|-marked, thus taking laryngeal relativism to the extreme, and killing it at the same time: if the representation is rigid, we are back to laryngeal realism, except reversed.

18 See, for example, the proposal of van der Hulst (1995) and his later attempts to provide structural configurations as phonological categories corresponding to phonetic substance. This structural view of melody, which finds an echo in the recent work of Pöchtrager & Kaye (2013), requires only one step towards 'substance-free' phonology, that is, assume arbitrary spell-out rather than one-to-one relations between particular structures and their phonetic interpretation.

References

Andersen, Hening 1969. The phonological status of the Russian labial fricatives. *Journal of Linguistics* 5: 121–127.

Arbaoui, Nora 2010. Les dix formes de l'arabe classique à l'interface syntaxe/phonologie. Ph.D. thesis, Université Paris 7.

Backley, Phillip 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Backley, Phillip & Kuniya Nasukawa 2009. Representing Labials and Velars: A single 'dark' element. *Phonological Studies* 12: 3–10.

Bendjaballah, Sabrina 1999. Trois figures de la structure interne des gabarits. Ph.D. dissertation, Université Paris 7.

Bendjaballah, Sabrina 2001. The negative preterite in Kabyle Berber. *Folia Linguistica* 34: 185–223.

Bendjaballah, Sabrina 2012. La grammaire des gabarits. Habilitation à Diriger des Recherches, University Paris 7.

Bendjaballah, Sabrina & Martin Haiden 2003. Templatic architecture. *Recherches Linguistiques de Vincennes* 32: 157–168.

Bendjaballah, Sabrina & Martin Haiden 2013. The representational anomalies of floating markers: Light prepositions in Taqbaylit of Chemini. *Challenges to Linearization*, edited by Theresa Biberauer & Ian Roberts, 331–376. Berlin: de Gruyter.

Bermúdez-Otero, Ricardo & Kersti Börjars 2006. Markedness in phonology and in syntax: The problem of grounding. *Linguistic Knowledge: Perspectives from Phonology and from Syntax*, edited by Patrick Honeybone & Ricardo Bermúdez-Otero, *Lingua* 116(5) (special issue): 710–756.

Boersma, Paul 1998. *Functional Phonology. Formalizing the Interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Bucci, Jonathan 2013a. Raddoppiamento Fonosintattico induit par l'accent et reduction vocalique en Italie: perspectives phonolgique et dialectologique. Ph.D. dissertation, Université de Nice, Sophia Antipolis.

Bucci, Jonathan 2013b. Voyelles longues virtuelles et reduction vocalique en coratin. *Canadian Journal of Linguistics* 58(3): 397–414.

Caratini, Emilie 2009. Vocalic and consonantal quantity in German: Synchronic and diachronic perspectives. Ph.D. dissertation, Nice University and Leipzig University.

Charette, Monik 2003. Empty and pseudo-empty categories. *Living on the Edge: 28 Papers in Honour of Jonathan Kaye*, edited by Stefan Ploch, 465–479. Berlin and New York: Mouton de Gruyter.

Chomsky, Noam 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam 1973. Conditions on transformations. *A Festschrift for Morris Halle*, edited by Stephen Anderson & Paul Kiparsky, 232–286. New York: Holt, Rinehart & Winston.

Chomsky, Noam 2000. Minimalist inquiries: The framework: Step by step. *Essays on Minimalist Syntax in Honor of Howard Lasnik*, edited by Roger Martin, David Michaels & Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam, Morris Halle & Fred Lukoff 1956. On accent and juncture in English. *For Roman Jakobson: Essays on the Occasion of His Sixtieth Birthday*, edited by Morris Halle, Horace Lunt, Hugh McLean & Cornelis van Schooneveld, 65–80. The Hague: Mouton.
Cyran, Eugeniusz 1996. Licensing properties of nuclei and principle ranking in Irish. *The Linguistic Review* 13: 1–31.
Cyran, Eugeniusz 2011. Laryngeal realism and laryngeal relativism: Two voicing systems in Polish? *Studies in Polish Linguistics* 6: 45–80.
Cyran, Eugeniusz 2014a. *Between Phonology and Phonetics: Polish Voicing*. Berlin: Mouton de Gruyter.
Cyran, Eugeniusz 2014b. The phonology and phonetics of obstruentization. *Crossing Phonetics-Phonology Lines*, edited by Eugeniusz Cyran & Jolanta Szpyra-Kozłowska, 31–46. Newcastle upon Tyne: Cambridge Scholars Publishing.
Cyran, Eugeniusz & Morgan Nilsson 1998. The Slavic [w > v] shift: A case for phonological strength. *Structure and Interpretation: Studies in Phonology*, edited by Eugeniusz Cyran, 89–100. Lublin: Folium.
D'Alessandro, Roberta & Tobias Scheer 2013. Phase head marking. *Linguistic Analysis* 38: 305–330.
D'Alessandro, Roberta & Tobias Scheer 2015. Modular PIC. *Linguistic Inquiry* 46: 593–624.
Embick, David & Rolf Noyer 2007. Distributed morphology and the syntax–morphology interface. *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand & Charles Reiss, 289–324. Oxford: Oxford University Press.
Faust, Noam 2013. Decomposing the feminine suffixes of modern Hebrew. *Morphology* 23: 409–440.
Guerssel, Mohand & Jean Lowenstamm 1990. The derivational morphology of the classical Arabic Verb. Ms., UQAM and University Paris 7.
Gussmann, Edmund 1998. Domains, relations and the English angma. *Structure and Interpretation. Studies in Phonology*, edited by Eugeniusz Cyran, 101–126. Lublin: Folium.
Gussmann, Edmund 2002. *Phonology: Analysis and Theory*. Cambridge: Cambridge University Press.
Gussmann, Edmund 2007. *The Phonology of Polish*. Oxford: Oxford University Press.
Hale, Mark & Charles Reiss 2000. Substance abuse and dysfunctionalism: Current trends in phonology. *Linguistic Inquiry* 31: 157–169.
Hale, Mark & Charles Reiss 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
Halle, Morris & Jean-Roger Vergnaud 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.
Hamann, Silke 2011. The phonetics-phonology interface. *The Continuum Companion to Phonology*, edited by Nancy Kula, Bert Botma & Kuniya Nasukawa, 202–224. London: Continuum.
Hamann, Silke 2014. Phonetics-phonology mismatches. Paper presented at Old World Conference in Phonology, Leiden 22–25 January.
Hamann, Silke & Paul Boersma (eds.) 2009. *Phonology in Perception*. Berlin: Mouton de Gruyter.
Harris, John 1990. Segmental complexity and phonological government. *Phonology* 7(2): 255–300.
Harris, John 1994. *English Sound Structure*. Oxford: Blackwell.
Harris, John 1996. Phonological output is redundancy-free and fully interpretable. *Current trends in Phonology. Model and Methods*, edited by Jacques Durand & Bernard Laks, 305–332. Salford, Manchester: ESRI.
Harris, John 2006. The phonology of being understood: Further arguments against sonority. *Lingua* 116(10): 1483–1494.
Harris, John 2009. Why final obstruent devoicing is weakening. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 9–45. Berlin and New York: Mouton de Gruyter.
Harris, John & Geoff Lindsey 1995. The elements of phonological representation. *Frontiers of Phonology. Atoms, Structures Derivations*, edited by Jacques Durand & Francis Katamba, 34–79. Harlow, Essex: Longman.
Hayes, Bruce 1984. The phonetics and phonology of Russian voicing assimilation. *Language Sound Structure*, edited by Mark Aronoff & Richard T. Oehrle, 318–328. Cambridge, MA: MIT Press.
Honeybone, Patric 2005. Diachronic evidence in segmental phonology: The case of laryngeal specifications. *The Internal Organization of Phonological Segments*, edited by Marc van Oostendorp & Jeroen van de Weijer, 319–354. Berlin: Mouton de Gruyter.

Hooper, Joan 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
Hulst, Harry van der 1995. Radical CV phonology: The categorical gesture. *Frontiers of Phonology. Atoms, Structures, Derivations*, edited by Jacques Durand & Francis Katamba, 80–116. Harlow, Essex: Longman.
Hulst, Harry van der 2014. The laryngeal class in RcvP and voice phenomena in Dutch. *Above and Beyond the Segments: Experimental Linguistics and Phonetics*, edited by Johanneke Caspers, Yiya Chen, Willemijn Heeren, Jos Pacilly, Niels O. Schiller & Ellen van Zanten, 323–349. Amsterdam: John Benjamins Publishing Company.
Hume, Elizabeth & Keith Johnson (eds.) 2001. *The Role of Speech Perception in Phonology*. New York: Academic Press.
Iverson, Gregory & Joseph C. Salmons 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12: 369–396.
Jensen, John 2000. Against ambisyllabicity. *Phonology* 17: 187–235.
Jensen, Sean 1994. Is ʔ an element? Towards a non-segmental phonology. *SOAS Working Papers in Linguistics and Phonetics* 4: 71–78.
Kaye, Jonathan 1989. *Phonology. A Cognitive View*. Hillsdale: Erlbaum. WEB.
Kaye, Jonathan 1990. 'Coda' licensing. *Phonology* 7: 301–330. WEB.
Kaye, Jonathan 1992. On the interaction of theories of Lexical Phonology and theories of phonological phenomena. *Phonologica 1988*, edited by Uli Dressler, Hans Luschützky, Oskar Pfeiffer & John Rennison, 141–155. Cambridge: Cambridge University Press. WEB.
Kaye, Jonathan 1995. Derivations and interfaces. *Frontiers of Phonology*, edited by Jacques Durand & Francis Katamba, 289–332. London & New York: Longman. Also in *SOAS Working Papers in Linguistics and Phonetics* 3, 1993, 90–126. WEB.
Kaye, Jonathan 2005. GP, I'll have to put your flat feet on the ground. *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*, edited by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster, 283–288. Berlin: Mouton de Gruyter.
Kaye, Jonathan, Jean Lowenstamm & Jean−Roger Vergnaud 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2: 305–328.
Kingston, John 2007. The phonetics-phonology interface. *The Cambridge Handbook of Phonology*, edited by Paul De Lacy, 435–456. Cambridge: Cambridge University Press.
Kiparsky, Paul 1982. From cyclic phonology to lexical phonology. *The Structure of Phonological Representations I*, edited by Harry van der Hulst & Norval Smith, 131–175. Dordrecht: Foris. WEB.
Kula, Nancy & Lutz Marten 2009. Defining initial strength in clusterless languages in Strict CV. *Strength Relations in Phonology*, edited by Kuniya Nasukawa & Phillip Backley, 251–284. Berlin: de Gruyter.
Lahrouchi, Mohamed 2001. Aspects morpho-phonologiques de la dérivation verbale en berbère (parler chleuh d'Agadir). Ph.D. dissertation, Université Paris 7.
Lampitelli, Nicola 2010. Nounness, gender, class and syntactic structures in Italian nouns. *Romance Languages and Linguistic Theory 2008*. Selected papers from 'Going Romance' Groningen 2008, edited by Reineke Bok-Bennema, Brigitte Kampers-Manhe & Bart Hollebrandse, 195–214. Amsterdam: Benjamins.
Lampitelli, Nicola 2011. Forme phonologique, exposants morphologiques et structures nominales: étude comparée de l'italien, du serbo-croate et du somali. Ph.D. dissertation, Université de Paris 7.
Lampitelli, Nicola 2013. The basic elements of inflection: Morphophonology of Bosnian nouns. *Formal Approaches to Slavic Linguistics 20: The Second MIT Meeting 2011*, edited by Alexander Podobryaev, 154–170. Ann Arbor: Michigan Slavic Publications.
Liljencrants, Johan & Bjorn Lindblom 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4): 839–862.
Lisker, Leigh & Arthur Abramson 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384–422.
Lowenstamm, Jean 1999. The beginning of the word. *Phonologica 1996*, edited by John Rennison & Klaus Kühnhammer, 153–166. La Hague: Holland Academic Graphics. WEB.

Lowenstamm, Jean 2003. *Grammaire et gabarits*. St.-Denis: Recherches Linguistiques de Vincennes 32.
Lowenstamm, Jean 2008. On little n, √, and types of nouns. *Sounds of Silence: Empty Elements in Syntax and Phonology*, edited by Jutta Hartmann, Veronika Hegedüs & Henk van Riemsdijk, 105–143. Amsterdam: Elsevier.
Lowenstamm, Jean 2011. The phonological pattern of phi-features in the perfective paradigm of noroccan Arabic. *Brill's Annual of Afroasiatic Languages and Linguistics* 3: 140–201.
Luo, Xiaoliang 2013. Vers une phonologie gabaritique du chinois. Analyse des interactions entre syllabe, ton et processus morphologiques. Ph.D. thesis, Université d'Orléans.
Marvin, Tatjana 2002. Topics in the stress and syntax of words. Ph.D. dissertation, MIT.
Mielke, Jeff 2008. *The Emergence of Distinctive Features*. Oxford: Oxford University Press.
Mołczanow, Janina 2008. *The Phonology of Glides in Russian*. München: Lincom Europa.
Nasukawa, Kuniya 2011. Representing phonology without precedence relations. *English Linguistics* 28: 278–300.
Nasukawa, Kuniya 2016. A precedence-free approach to (de-)palatalisation in Japanese. *Glossa* 1: article 9, DOI: http://doi.org/10.5334/gjgl.26.
Nespor, Marina & Irene Vogel 1986. *Prosodic Phonology*. Dordrecht: Foris.
Oostendorp, Marc van 1994. Affixation and integrity of syllable structure in Dutch. *Linguistics in the Netherlands 1994*, edited by Reineke Bok-Bennema & Crit Cremers, 151–162. Amsterdam: Benjamins.
Pagliano, Claudine 2003. L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie. Ph.D. dissertation, Université de Nice.
Pak, Majorie 2008. The postsyntactic derivation and its phonological reflexes. Ph.D. dissertation, University of Pennsylvania.
Passino, Diana 2008. *Aspects of Consonantal Lengthening in Italian*. Padova: Unipress.
Passino, Diana 2011. Phonological traces of morphosyntactic phases: The phonology/syntax interface in the varieties of Italian: I luoghi della traduzione. *Atti del XLIII Congresso Internazionale della Società di Linguistica Italiana*, edited by Giovanna Massariello & Serena dal Maso, 849–864. Rome: Bulzoni.
Passino, Diana 2014a. Dialectes italo-romans et théorie phonologique: un rapport d'enrichissement mutuel. Habilitation Thesis, Université de Nice Sophia Antipolis.
Passino, Diana 2014b. An element-based analysis of Italian nominal inflection. *Selected Proceedings of Décembrettes 6, Morphology in Bordeaux*, edited by Fabio Montermini, Gilles Boyé & Jesse Tseng, 63–75. Somerville, MA: Cascadilla.
Pöchtrager, Markus & Jonathan Kaye 2013. GP 2.0. *SOAS Working Papers in Linguistics* 16: 51–64.
Prince, Alan 1985. Improving tree theory. *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, edited by Mary Niepokuj, Mary VanClay, Vassiliki Nikiforidou & Deborah Feder, 471–490. Berkeley: University of California, Berkeley.
Prunet, Jean-François 1986. Spreading and locality domains in phonology. Ph.D. dissertation, McGill University at Montreal. Published in 1992 by Garland Press.
Prunet, Jean-François 1987. Liaison and nasalization in French. *Studies in Romance Languages*, edited by Carol Neidle & Rafael Nuñez Cedeño, 225–235. Dordrecht: Foris.
Reiss, Charles 2007. Modularity in the 'sound' domain: Implications for the purview of universal grammar. *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand & Charles Reiss, 53–79. Oxford: Oxford University Press.
Riemsdijk, Henk van 1978. *A Case Study in Syntactic Markedness*. Lisse: Peter de Ridder.
Rubach, Jerzy 1996. Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27: 69–110.
Rucart, Pierre 2006. Morphologie gabaritique et interface phonosyntaxique. Aspects de la morphologie verbale en afar. Ph.D. dissertation, Université Paris 7.
Samuels, Bridget 2009. The structure of phonological theory. Ph.D. dissertation, Harvard University.
Scheer, Tobias 1999. A theory of consonantal interaction. *Folia Linguistica* 32: 201–237. WEB.
Scheer, Tobias 2000. De la Localité, de la Morphologie et de la Phonologie en Phonologie. Habilitation thesis, Université de Nice. WEB.

Scheer, Tobias 2004. *A Lateral Theory of Phonology: Vol.1: What Is CVCV, and Why Should It Be?* Berlin: Mouton de Gruyter.

Scheer, Tobias 2008a. Spell out your sister! *Proceedings of the 27th West Coast Conference on Formal Linguistics*, edited by Natasha Abner & Jason Bishop, 379–387. Somerville: Cascadilla. WEB.

Scheer, Tobias 2008b. Why the prosodic hierarchy is a diacritic and why the interface must be direct. *Sounds of Silence: Empty Elements in Syntax and Phonology*, edited by Jutta Hartmann, Veronika Hegedüs & Henk van Riemsdijk, 145–192. Amsterdam: Elsevier. WEB.

Scheer, Tobias 2009a. External sandhi: What the initial CV is initial of. *Studi e Saggi Linguistici* 47: 43–82. WEB.

Scheer, Tobias 2009b. Representational and procedural sandhi killers: Siagnostics, distribution, behaviour. *Czech in Formal Grammar*, edited by Mojmír Dočekal & Markéta Ziková, 155–174. München: Lincom. WEB.

Scheer, Tobias 2011. *A Guide to Morphosyntax-Phonology Interface Theories. How Extra-Phonological Information Is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.

Scheer, Tobias 2012a. At the right edge of words (again). *Festschrift for Glyne Piggott: McGill Working Papers in Linguistics* 22: 1–29.

Scheer, Tobias 2012b. *Direct Interface and One-Channel Translation: A Non-Diacritic Theory of the Morphosyntax-Phonology Interface: Vol.2 of a Lateral Theory of phonology*. Berlin: de Gruyter.

Scheer, Tobias 2014a. The initial CV: Herald of a non-diacritic interface theory. *The Form of Structure, the Structure of Form: Essays in Honor of Jean Lowenstamm*, edited by Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi & Nicola Lampitelli, 315–330. Amsterdam: Benjamins.

Scheer, Tobias 2014b. Spell-out, post-phonological. *Crossing Phonetics-Phonology Lines*, edited by Eugeniusz Cyran & Jolanta Szpyra-Kozłowska, 255–275. Newcastle upon Tyne: Cambridge Scholars Publishing.

Scheer, Tobias 2016a. Melody-free syntax and phonologically conditioned allomorphy. *Morphology* 26: 341–378.

Scheer, Tobias 2016b. Juncture and Boundary. *Oxford Bibliographies Online* publication at www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0074.xml.

Schwartz, Jean-Luc, Luis-Jean Boë & Christian Abry 2007. Linking dispersion-focalization theory and the maximum utilization of the available distinctive features principle in a perception-for-action-control theory. *Experimental Approaches to Phonology*, edited by Maria-Josep Solé, Patrice Beddor & Manjari Ohala, 104–124. Oxford: Oxford University Press.

Ségéral, Philippe 2000. Théorie de l'apophonie et organisation des schèmes en sémitique. *Research in Afroasiatic Grammar*, edited by Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky, 263–299. Amsterdam & Philadelphia: Benjamins. WEB.

Ségéral, Philippe & Tobias Scheer 2001. Abstractness in phonology: The case of virtual geminates. *Constraints and Preferences*, edited by Katarzyna Dziubalska-Kołaczyk, 311–337. Berlin and New York: Mouton de Gruyter.

Ségéral, Philippe & Tobias Scheer 2008. The Coda Mirror, stress and positional parameters. *Lenition and Fortition*, edited by Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral, 483–518. Berlin: Mouton de Gruyter. WEB.

Starke, Michal 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36: 1–6.

Stevens, Kenneth 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. *Human Communication: A Unified View*, edited by Peter B. Denes & Edward E. David Jr., 51–66. New York: McGraw Hill.

Uriagereka, Juan 2011. Derivational cycles. *The Oxford Handbook of Minimalism*, edited by Cedric Boeckx, 239–259. Oxford: Oxford University Press.

Zwicky, Arnold & Geoffrey Pullum 1986. The principle of phonology-free syntax: Introductory remarks. *Ohio State University Working Papers in Linguistics* 32: 63–91.

12 Dependency Phonology

Harry van der Hulst and Jeroen van de Weijer

12.1 Introduction

Whenever two minimal units enter into a relation, they form a construction and, typically, the relation between units in a construction will not be equal; it is asymmetrical. This is, in short, the heart of the wisdom that Dependency Phonology (DP), or Dependency Grammar more broadly, has contributed to linguistic theory. In contrast with a constituency-based approach, there are no constituents, no 'consist of' relations, in the dependency approach. In language, asymmetrical relations are found everywhere where two units combine: in stress languages, two syllables are joined into a 'foot', where one will be stressed and the other unstressed. In morphology, two nouns can form a compound with one being semantically dominant as well as, typically, determining the word class. In syntax, one word in a phrase will function as the syntactic 'head'. Even in single segments such as affricates there is an asymmetric relation between the phonetic parts of the segment. The status and implementation of this head-dependency relation (HDR) in both segmental and suprasegmental structure is the defining feature of the DP framework, which we will discuss in this contribution.

The organization of this chapter is as follows. Section 12.2 discusses the basic principles of the DP approach. Section 12.3 reviews some proposals for revision or further extensions of the DP model that have been made in the literature.[1] While these revisions mostly focus on the structure of segments, section 12.4 discusses suprasegmental structure, starting with the notion of syllable structure and then moving on to the distinction between word and utterance structure. Section 12.5 deals with the manner in which DP allows the expression of phonological alternations. Section 12.6 compares DP to other phonological models, and section 12.7 offers a brief conclusion.[2]

As a preamble, a note on the term 'dependency'. This term has been used in a variety of ways, as also noted by Ewen (1995). In Feature Geometry proposals (Clements 1985; Sagey 1986), the term can refer to the hierarchical relation between a mother node and its daughter(s), i.e. as the inverse of dominance: no headedness in the DP sense is assumed; this is the sense in which McCarthy (1988) (and others) have used the term. Mester (1986, 1988) allows different features, residing on different tiers, to be dependent on each other, such that spreading one 'drags along' the other. A related concept is 'government' (the inverse of dependency), as in Government Phonology (Kaye et al. 1985, 1990) (see Chapters 9–11 in this book, as well as section 12.6 below).

12.2 Basic principles

12.2.1 Dependency and structural analogy

DP adopts the basic premise of Dependency Grammar, which is that linguistic units enter into *constructions* that are characterized by a relation of dependency between heads and dependents. The relation of dependency is applied in both the *plane* that combines meaningful (conceptually based) basic units into larger constructs (i.e. syntax; the *content plane*) and in the phonological plane (whose constructs involve meaningless, perceptually based basic units: the *expression plane*).[3] Fundamental to Anderson's work is the *Structural Analogy Assumption* (see also Anderson 1971; Anderson 1987; Anderson 2004; Anderson 2011a, 2011b, 2011c; Bauer 1994; and Staun 1996a for discussion), which holds that structural relations and principles are the same in both planes of grammar. The planes therefore primarily differ in terms of the sets of their basic units, i.e. their alphabets, which are determined by the interface with phonetic substance (for the expression plane) and conceptual meaning (for the content plane).[4] The assumption of structural analogy has roots in Louis Hjelmslev's theory of glossematics (e.g. Hjelmslev 1953). It might seem that this assumption runs counter to the modularity assumption that is prevalent in Generative Grammar (and Cognitive Science in general), but this is only true if we assume that recognizing different modules (of grammar or of the mind) somehow entails that these modules must have radically different organizations. Anderson, as do we, adopts the more plausible assumption that different modules follow the *same* principles of organization to the extent that this is possible. Indeed, there is no reason to believe that the notion of dependency, or any of the other basic principles that we will discuss, are limited to grammatical modules.[5] By taking analogies between the two planes as non-accidental and in fact reflecting the relevance of general principles in both domains, Anderson's Dependency Grammar takes a stance that has obvious implications for the debate about an alleged Universal Grammar that merely comprises a syntactic system, relegating phonology to a separate 'expression system' (e.g. Hauser et al. 2002). We will follow Anderson in claiming that the existence of profound analogies between the expression plane and the content plane strongly argues against separating the cognitive systems that permit humans to construct a mental grammar for their language(s) in this radical fashion. At the same time, we agree with Anderson that there is little reason to believe that these analogies reflect principles that are confined to an alleged innate Universal Grammar, however construed.

Dependency structures form an alternative to constituency-based approaches: there is a principled distinction between the two. In a dependency approach, all nodes are associated to units from the alphabet. This means that there are no phrasal nodes that dominate non-terminal nodes. This fundamental difference may be obscured by several factors, however. Firstly, constituent structure in Generative Grammar has been augmented with the notion of headedness ever since Chomsky (1980). Constituents are said to be headed, with the head being a basic, i.e. lexical, unit that determines the characteristic properties of the phrase it heads. The resulting hybrid approach (constituency-cum-headedness) has also found its way into Generative Phonology (specifically in theories of suprasegmental structure). Secondly, depending on how dependency graphs are conceived, it is often very easy to map a dependency graph onto a more familiar-looking constituent structure, especially when the relationship of subjunction is used (see section 12.4). While such a mapping may be deemed to serve no purpose, it is nonetheless the case that the resemblance may obscure the principled difference. Despite these factors that might blur the distinction to the casual observer, the rejection of constituent structure is fundamental to Dependency Grammar.

Anderson makes a distinction between two types of dependents: complements (dependents that the head requires) and adjuncts (optional modifiers of the head). We will illustrate this distinction in section 12.4, where we discuss the DP approach to syllable structure. The dependency approach that is reviewed in this chapter has been developed by John Anderson (and a number of other phonologists) over the last five decades. With reference to the alphabets for each plane, Anderson has advocated a strong substantive, or grounded, position. Phonological units and structures are firmly grounded in perceptual acoustics, while the basic units and structures of morphosyntax are grounded in meaning/conceptual structure. Groundedness also extends to structure, i.e. the formation of constructions, in both planes. Headedness in both planes correlates with a substantive notion of *cognitive salience*. The substance-based approach stands in stark contrast to so-called substance-free theories (see e.g. Hale and Reiss 2000; Blaho 2008).

We add a word about the 'sociology' of Dependency Grammar here. While an appeal to dependency as the organizational relation that binds words together into sentences has deep roots in ancient approaches to language (Percival 1990), it is due to the work of a few scholars that this approach has developed into a branch of linguistics in modern times. In particular, Tesnière (1959) is a foundational work, but other relevant references are Hays (1964); Gaifman (1965); Heringer (1967); and Marcus (1967). We refer to Anderson and Ewen (1980); Anderson and Durand (1987); van der Hulst (2006); and of course Anderson and Ewen (1987) for general overviews of the dependency approach to phonology. As far as we can tell, Anderson is the only linguist who has applied this approach to phonology. While, as we will show, various ingredients of his proposal (developed in the early 1970s, in collaboration with others) bear strong resemblances to versions of Generative Phonology that were developed in the 1970s and 1980s, these later developments took place independently, mostly in the United States. Indeed, Anderson, working in Edinburgh (Scotland) did not 'found a school' which could exercise influence in other countries, let alone continents. We are aware of only one dissertation in this framework that was written in the United States (Kang 1991). DP's major resource remains *Principles of Dependency Phonology* (Anderson and Ewen 1987). Various other phonologists have also made contributions to DP, mostly with publications in European journals and in some edited volumes.[6]

12.2.2 Segmental structure: monovalency, grouping, dependency and contrastivity

In this section, we focus on segmental structure. In the segmental domain, DP introduced at least six important innovations, several of which date back to early publications by John Anderson and Charles Jones (Anderson and Jones 1972, 1974):[7]

(1) Segmental structure
- Phonological primes (called 'components') are monovalent
- Phonological primes are organized into intrasegmental classes (called 'gestures')
- Combinations of primes and of classes enter into a head-dependency relationship
- The same phonological primes figure in the representation of vowels and consonants
- Representations are minimally specified
- Some primes may occur in more than one class

We must note that these aspects are largely independent and, as such, may be shared (in part) with other approaches (see section 12.6). The following sections deal with specific characteristic topics in DP: monovalency (2.2.1), the idea that vowel structure is organized in a triangular way (12.2.2.2), segment-internal grouping (12.2.2.3) and minimal specification (12.2.2.4).

12.2.2.1 Monovalency

With little if any precedent in phonology, Anderson and Jones (1972, 1974) proposed, in response to the tradition of binary features (Jakobson et al. 1952; Chomsky and Halle 1968) that the basic building blocks of phonology are monovalent (i.e. have only one value): they are unary instead of binary.[8] While DP uses the term *component*, we will here, following Government Phonology (Kaye et al. 1985), refer to these unary features as *elements*.[9]

An important distinction between the binary and unary approaches is the fact that the binary approach allows reference to both values of a distinctive feature. For example, in the case of the feature [±voice], binary theories recognize both a class of voiced and a class of voiceless segments, whereas unary approaches only allow reference to the class that is positively specified with an element. (That is, if we disallow reference to the absence of a property in a unary model.) Given this fact, a unary approach should count as the null hypothesis because it is more restrictive, placing the burden of proof on proponents of binary features; see Kaye (1988). Historically, features entered the phonological arena as binary units (see again Jakobson et al. 1952; Chomsky and Halle 1968) and for this reason it is often assumed that unarists have to defend their position against the binary approach. However, from a methodological point of view, once a contrast has been established, the initial hypothesis must be that opposition is encoded in monovalent terms, thus claiming that 'the other value' is a phonological non-entity. This hypothesis can be falsified either by facts that require reference to the other pole (still privative), or by facts that require reference to both poles. Facts of the latter type necessitate an equipollent characterization of the opposition, either in terms of a binary feature or in terms of two unary features.

Apart from the fact that a unary feature theory is more restrictive, Anderson and Jones also motivate their proposal on the argument that binary features present a problem for the notion of markedness. This had in fact also been noticed by Chomsky and Halle (1968), who devoted a 'late chapter' (chapter 9) in their *Sound Pattern of English* (SPE) to the fact that a theory using binary features cannot cope with certain recurrent asymmetries between the two values of some, or perhaps all, features. Comparing the vowels /ü/ and /i/, they note, as others did before them, that the roundness of /ü/ and the non-roundness of /i/ should be weighted differently, in that front vowels, in the absence of a rounding contrast, are always [−round]. Another indication of the asymmetry comes from cases of neutralization. For example, in the domain of obstruents, where voicing is typically distinctive, voiced obstruents seem more restricted in that, if the opposition is neutralized word-finally, the voiceless obstruents emerge. Unary features allow for a direct and, in fact, literal expression of markedness. The vowel /ü/ is more marked than /i/ because it must bear the mark of roundness, both vowels being specified as front. Likewise, voiced obstruents are more marked than voiceless obstruents (at least in most contexts; see following discussion), since they bear an element corresponding to [+voice] and voiceless obstruents do not.

In binary feature theories, the most straightforward expression of the asymmetry between the two values is to leave the 'expected' values literally unmarked. (Hence these values themselves became known as 'unmarked values'). Thus, the unmarked value of [round] (for front vowels) is minus, and the unmarked value for voice (in obstruents) is also minus.[10]

This approach is referred to as Underspecification Theory (Halle 1959, et seq.). However, for technical reasons Chomsky and Halle (1968) could not appeal to underspecification (see Stanley 1967), but instead adopted special *m/u* values for features (alongside the plusses and minuses) and a set of markedness (and linking) conventions (see Kean 1975; van Lessen Kloeke 1982). This theory of markedness, however, was soon abandoned, and eventually underspecification made a comeback (Ringen 1978; Kiparsky 1982; Archangeli 1984). Kiparsky and Archangeli proposed that unmarked values should not only be unspecified if they are redundant (i.e. in the absence of a contrast) but also when contrast is in place. This approach, which encodes unmarkedness in terms of non-specification, came to be known as Radical Underspecification Theory.[11]

On one view, a monovalent approach represents an extreme form of radical underspecification. The claim is simply that unmarked or default values play no role in the phonology whatsoever. However, we must note that the issue of using under- or non-specification is not confined to binary feature systems: it is also relevant in monovalent theories (see e.g. Durand 1988 and section 12.2.2.4 below).

Clearly, while a single-valued system reflects the spirit of (radical) underspecification by establishing a direct correlation between markedness and complexity, it does so in a more rigorous way. Despite the fact that radical underspecification theories ban one value from phonological representations, the 'unmarked' default one, the option is left open that these values are filled in at some point in the derivation, after which they may start playing a role in the phonology by appearing in rules as targets, changes or environments. More dramatically, it has been argued that the markedness of a value may not be universal in that some languages may show a 'markedness reversal' (see e.g. Battistella 1990; de Lacy 2006). This, then, allows for a situation in which [+voice] is the default value for (e.g. final) obstruents in some language. Monovalent theories do not allow for markedness reversals, nor do they allow the 'unmarked value' to become active in the phonology. The 'unmarked value' is a phonological non-entity.

The reader might ask how, if this is the case, markedness can ever be contextual. That is, how can we account for the fact that [−round] is unmarked for front vowels, requiring the specification of [+round] for front rounded vowels, among back non-low vowels, [+round] is the unmarked value, which would suggest that [−round] must be specified for back non-round vowels if there is a contrast? A unary system that uses the unary features [front] and [round] would seem to be committed to representing the 'less marked' /u/ as more complex than the more marked /ɯ/:

(2)	/i/	/ü/	/ɯ/	/u/
	front	front	–	–
		round		round

We will return to this conundrum below, which has haunted unary systems for a long time.

All things being equal, a unary approach is more restrictive than a binary approach. However, in practice, when comparing different feature theories, all things are never equal. Theories can differ in terms of which specific features they have, what kinds of intrasegmental relations (such as head-dependency) are used, and what kinds of formal manipulations ('rules') they permit. The issue of fair comparison becomes even more complicated when monovalent approaches include primes that seem to be polar opposites. We see this in some non-DP models that use unary features, for example when two monovalent feature [ATR] and [RTR] are proposed (see Steriade 1995 and others). Van de Weijer (1992, 1993,

1996) proposed the opposite manner features [stop] and [cont], with the idea that both define recurrent natural classes. Van der Hulst (2005, in prep.) argues for a particular approach that makes systematic use of primes that form pairs of polar opposites (see section 12.3.4). Adopting apparently polar opposites is not equivalent to adopting a binary feature, however. Under usual assumptions, two values of a binary feature cannot be combined within a segment, or if they can, this must lead to phonetic sequencing (as in [−cont][+cont] proposals for affricates). Unary features, on the other hand, even when apparently opposites, may be combined to represent an intermediate category. This will be illustrated in section 12.2.2.2, where we will discuss the specific DP proposals for unary feature sets that have been proposed within DP. This will also introduce the notion of intrasegmental dependency.

12.2.2.2 The triangular set

Moving beyond the issue of the 'arity' of features, we will now discuss the specific set of elements that have been proposed in DP. Anderson and Jones (1972, 1974) focused on the representation of vowels. Given this limitation, this early publication did not propose a 'complete' set of phonological elements and therefore did not develop the notion of grouping elements into subsegmental units (classes, gestures). They introduced the characteristic and basic |a|, |i|, |u|[12] set, showing how these units can be used to represent vowels, allowing them to occur by themselves or in combinations. Let us take a closer look at the DP proposal for vowel representation.[13] Clearly, the DP system differs from the SPE system not only by using unary rather than binary features, but also by choosing different phonetic parameters for characterizing the vowel space. Whereas the SPE system is bidirectional (just like, for instance, the unary feature system proposed by Sanders 1972), since it only uses the high-low and the front-back dimensions in the description of vowels, lip rounding being superimposed on these two dimensions, the feature system of DP is tridirectional.[14]

Characteristic of tridirectional feature systems is the fact that they employ at least three basic primes in their element set, corresponding to the three corners of the vowel triangle. In DP, these elements are first and foremost grounded in acoustic percepts. The three basic primes are commonly represented by the symbols |i|, |u| and |a|, after the vowels that these elements represent if they occur alone.

(3) The basic primes of tridirectional unary feature systems for vowels:

	Acoustic	Articulatory
|i|	acuteness/sharpness	frontness
|u|	gravity/flatness	roundness
|a|	sonority	lowness

From a phonetic point of view, these elements are clearly basic. They constitute the so-called quantal vowels (Stevens 1972), that is, they are the acoustically most stable vowels, in that their acoustic effects can be produced with a fairly wide range of articulatory configurations. In addition, these three vowels are maximally distinctive, both from an acoustic and an articulatory point of view (see Liljencrants and Lindblom 1972 and related work). Moreover, /i/, /u/ and /a/ are also basic as far as phonology is concerned. They constitute the canonical three-vowel system, and they typically are also the first vowels that children acquire. The choice of |i|, |u| and |a| as basic vocalic elements is therefore well-motivated, both phonetically and phonologically.

With the aid of these three vowel elements, at most seven vowels can be characterized, if we bear in mind that they can be used not only in isolation, but also in combination with each other:

(4) {|i|}[15] {|u,i|} {|u|}
{|i,a|} {|u,i,a|} {|u,a|}
{|a|}

It will be obvious that these seven representations do not exhaust the maximal number of different vowels that are found in the language systems of the world, nor, more crucially, possibly richer (or simply different) sets of vowels that occur in specific languages. To express vowel systems containing nine or even more vowels, additional ways are needed to represent the total number of vowels in terms of (combinations) of the three basic vocalic elements. In principle, there are two ways in which this increase of the combinatorial potential of the three features could be achieved. Features might either occur more than once in a particular representation, or one of the features in a feature combination might be prominent relative to another feature (or features). Of these two conceivable positions, the former is defended by Schane (1984) (in Particle Phonology (PP); see section 12.6), while DP (as well as Government Phonology; see section 12.6) invokes the concept of dependency to arrive at a larger number of possible representations.

Compare, for instance, the DP and PP representations of the vowel /ɛ/ in the partial vowel system in (5):

(5) /i/ DP: {|i|} PP: {I}
/e/ DP: {|i;a|} PP: {IA}
/ɛ/ DP: {|a;i|} PP: {IAA}
/a/ DP: {|a|} PP: {A}

Here dependency is expressed using the symbol ';', {A;B} being read as 'B is dependent on A', or 'A governs B'; see (6) for another notation.[16]

As shown, in DP, elements are not just joined in a simple, symmetrical combination, but they can also enter into a relationship in which one element is relatively prominent, i.e. the 'head' and the other element is the dependent. If a language has just one mid-series, the dependency relation can remain unspecified. We note at this point that it is commonly assumed in phonology that contrastive use of phonetic properties involves a binary opposition, which can be expressed with a binary feature or a unary feature (presence vs. absence). Apparently gradual differences along a phonetic dimension can be represented with more than one feature. This can be seen in binary systems where two or more features that refer to height or aperture jointly capture a three- or four-level height distinction. In DP, such gradual effects are captured by invoking combination of elements and their various dependencies. With reference to sonority we will discuss this in section 3.4.

In addition, two elements can even entertain a relation in which neither feature is dominant, a relationship which DP calls 'mutual/bilateral dependency'. Thus we arrive at the set of dependency relationships in (6), in two alternative notations that Anderson and Ewen

(1987) use to express dependency; the braces stand for “a class of segments characterized by the element structure in question” (p. 151).

(6) a. {|X;Y|} or {|X⇒Y|} – Y is dependent on X
b. {|Y;X|} or {|Y⇒X|} – X is dependent on Y
c. {|X:Y|} or {|X⇔Y|} – X and Y are mutually dependent

By allowing the features to enter into a relationship of ‘mutual dependency’ with |a|, a relationship in which neither element counts as the head, DP maximally generates the following set of representations on the basis of the features |i|, |u| and |a|:

(7) The maximum number of combinations of |i|, |u| and |a| in DP:

{|i|}	{|u,i|}	{|u|}
{|i;a|}	{|u,i;a|}	{|u;a|}
{|i:a|}	{|u,i:a|}	{|u:a|}
{|a;i|}	{|a;u,i|}	{|a;u|}
	{|a|}	

Implicitly, it is assumed that {|i,a;u|}, {|u;a,i|}, {|i;a,u|}, {|a,u;i|} do not result in phonetically distinct vowels, i.e. that they result in phonetically equivalent events. This means that the combination |u,i| seems to behave like a unit, such that |u| and |i| cannot occur on opposite ends of the dependency relation. In other words, this combination of elements does not seem to show a dependency asymmetry.

Although the system of DP would in principle allow for the gradual oppositions {|i|} vs. {|i;u|} vs. {|i:u|} vs. {|u;i|} vs. {|u|}, it turns out, as Anderson and Ewen (1987: 275) observe, that “in virtually all languages, we find at each height maximally one segment containing both |i| and |u|; in other words, dependency relationships holding between |i| and |u| are not required”.[17] Yet, although they may not be required in practice, the fact remains that nothing in the theoretical framework of DP renders dependency relations between the features |i| and |u| impossible on a principled basis. Van der Hulst (2005, in prep.) proposes to use these two possible ways of combining the color elements to represent the two kinds of rounded vowels in Swedish (e.g. Riad 2014).

Staying with the DP proposal to not allow |i| and |u| to combine in two ways, at most eight front vowels and four back vowels can be represented, plus the low vowel. This is still, however, not enough to characterize all possible vowels and vowel systems in the world’s languages. In particular, the central vowels and/or the back unrounded vowels cannot be represented on the basis of (6) alone. Here the ‘and/or’ refers to the fact that it is not certain that central and back unrounded are distinct phonological categories, although the former class, according the IPA-system, allows both rounded and unrounded vowels. The mid rounded vowels perhaps require a separate class in any event. This brings us back to the issue raised in (2) of the representation of /u/ vs. /ɯ/, which raised the question how this contrast can be represented without running into a ‘markedness paradox’. To solve this problem, there have been various proposals to separate backness from roundness, thus ‘splitting up’ the U-element.

Van der Hulst (1988) addresses this issue in the context of a specific proposal that builds on the fact that elements in head position contribute more strongly to the resulting vowel than the same element in dependent position; indeed such elements have greater perceptual and thus cognitive salience. This means that phonetic interpretation is sensitive to the head or dependent status of an element. Van der Hulst (1988) pushes this one step further by proposing that a specific phonetic interpretation of elements correlates with their head and dependent occurrence, as summed up in (8), using articulatory rather than acoustic labels:[18]

(8)

Interpretation of |u|	Head:	Velar constriction
	|	
	Dependent:	Rounding
Interpretation of |i|	Head:	Palatal constriction
	|	
	Dependent:	Advanced Tongue Root
Interpretation of |a|	Head:	Pharyngeal constriction
	|	
	Dependent:	Openness

This proposal allows an element to occur twice, which is not a standard assumption in DP. By itself, however, this system does not solve the markedness issue, since /u/ is more complex than /ɯ/ in (9):

(9)

/i/	/ü/	/ɯ/	/u/
i	i	u	u
u	u		

We will return to van der Hulst's proposal (which was developed in van der Hulst 1989 and also in work by Norval Smith and students; e.g. Botma 2004, 2009; Botma and Smith 2006, 2007; Smith 1988) in section 12.3, where we will discuss the idea of using an element more than once in a representation.[19] Here we will focus on the overt recognition of the dual character of |u|, which has also been acknowledged in other proposals. A number of phonologists, notably Lass (1984) and Rennison (1986), have argued that these two aspects of |u| should in fact be given independent status, thus splitting up |u| into two features, |ω| ('labiality' or 'roundness') and |ɯ| ('velarity' or 'high backness'), which still entails the same problem as in (9): /u/ comes out as more marked:

(10)

/i/	/ü/	/ɯ/	/u/
i	i	ɯ	ɯ
	ω	ω	

However, these various proposals do not solve the problem of how to represent central vowels. To deal with the problem of central vowels, Anderson and Ewen (1987) propose a different solution. To the vowel /ɯ/ they assign not only the two color elements, but also a new element: |ə|, the centrality element:

(11) The representation of /u/: standard DP: {|u|} The representation of /ɯ/: standard DP: {|u,i,ə|}

While this proposal solves the markedness asymmetry by representing central vowels as more complex, another solution that could be considered is to represent /ɯ/ as devoid of any elements; this is in fact what Anderson (2011c) suggests. The idea that one vowel can be represented as the null set has other precedents, especially with regard to one of the central vowels, in particular the schwa (see e.g. S. Anderson 1982).[20] At first sight, this makes this vowel the least complex, but if we limit the markedness-complexity correlation to segments that are positively specified, we can add the special clause that a segment that is devoid of any property is the most 'marked' vowel, due to the fact that it misses any perceptual salience, which is worse than mixing two perceptual images, as in vowels that combine two or more elements.[21] The proposal to acknowledge the 'null option' (lacking elements) may obviate the need for the centrality element, although it is not clear how central vowels of different heights will be represented, if the null 'element' is not allowed to combine. Van der Hulst (in prep.) solves this problem by introducing a fourth element, |∀|, similar to the centrality element, which can enter in to a dependency relationship with the element |A| to represent four colorless vowels of different heights:

(12)

	I	IU	Placeless	UI	U
∀	i ~ ɪ	y/Y	ɨ ~ ɯ	ʉ	u/ʊ
∀A	e	ø	ɘ ~ ɤ	ɵ	o
A∀	ɛ	œ	ɜ ~ ʌ	ɞ	ɔ
A	æ	ɶ	a	ɑ	ɒ

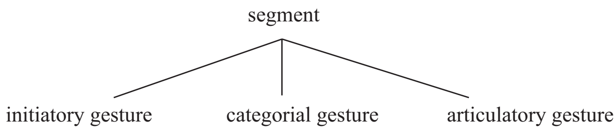
This chart also contains two series of non-back round vowels, based on the headedness of combinations of the two color elements. It does not distinguish between advanced and non-advanced vowels (as indicated for the high series, which requires an element for the expression of tongue root position), which we will discuss in the next section, after first introducing the notion of grouping.

12.2.2.3 Grouping

The relevance of feature grouping has long been recognized in DP. While it was not part of the original proposal by Anderson and Jones (1974), Lass and Anderson (1975) and Lass (1976) offer a number of specific arguments that support the view that the matrix characterizing the segment should be split up into at least two submatrices, or *gestures*. This subdivision into element sets reflects the fact that phonological processes can refer precisely (e.g. delete or spread) to either of these gestures, the other gesture being unaffected (cf. the so-called 'stability effects' of Autosegmental Phonology; Goldsmith 1976). Lass (1976) discusses cases of reduction of full consonants to the glottal consonants [h] and glottal stop [ʔ], which occur, for instance, in many varieties of Scots (cf. also Lass 1984: 113–115), which show the independence of the laryngeal features vis-à-vis the oral features, a proposal also made in Thráinsson (1978) on the basis of Icelandic preaspiration data and subsequently in various versions of Feature Geometry. The DP arguments for grouping are essentially analogous to the arguments that have been presented for feature classes in Feature Geometry (see Clements 1985; Sagey 1986).

In early DP work, the bipartite division that was suggested by Lass and Anderson (1975) into a laryngeal gesture and an oral gesture was replaced by the following proposal for a tripartite gestural division of segments (Anderson and Ewen 1980; Ewen 1980; Lass 1984), by splitting the oral gesture into a gesture for major class and manner-like distinctions (the categorial gesture), and a strictly articulatory (place) gesture. The term 'gesture' here is used completely equivalently to the way in which 'class node' is used in Feature Geometry, where one segment (the unity of which is expressed by the root node) consists of various class nodes.

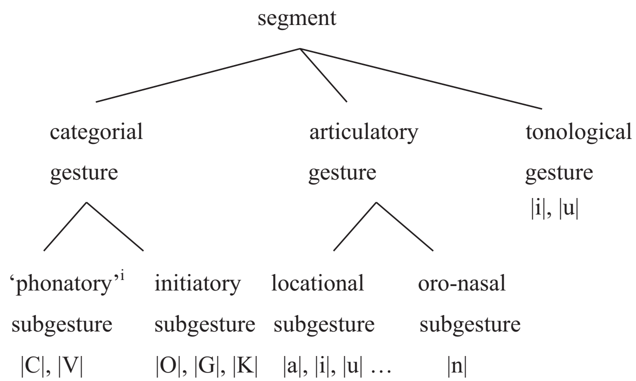
(13)



The initiatory gesture contains elements expressing airstream properties and glottal states.

Ewen (1986: 205) extends this model by recognizing two major 'super' gestures, the categorial and the articulatory gesture, both of which contain two subgestures. The categorial gesture contains a 'phonatory' subgesture (for elements expressing manner or stricture properties and major class distinctions) and the initiatory subgesture (as before, for airstream properties and glottal states). The articulatory gesture contains the locational subgesture (with elements for place properties) and an oro-nasal subgesture containing just one element (viz. nasal). In addition, a tonological gesture is added:

(14)



[i] We put this term between quote marks because the use of the term 'phonatory' here is unfortunate; it essentially is about major class and manner.

The locational elements listed in (14) are not an exhaustive set; see below.

We will discuss the structure displayed in (14) in more detail, following Anderson and Ewen (1987) (henceforth AE). The proposals which AE make for the tonological gesture are sketchy (see also p. 238–9). Most work focuses on the development of the 'phonatory' subgesture (for manner and major class distinctions) and the locational gesture (for place). We will discuss these two subgestures in turn.

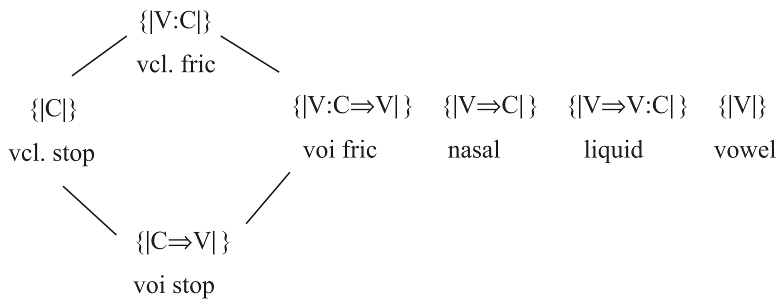
The 'phonatory' subgesture contains two elements, |V| and |C|, which AE define as follows: "|V|, a component which can be defined as 'relatively periodic', and |C|, a component of 'periodic energy reduction'" (p. 151). As mentioned above, from the start DP adopted the view that the primary interpretation of elements is acoustic, a position that Government Phonology has adopted as well. They then continue:

> |V| and |C| differ from the [Jakobsonian] vocalic and consonantal distinctive features in that the presence of, say, |V| in a segment does not necessarily imply that the segment is in a simple binary opposition to an otherwise identical segment not containing |V|. Rather [. . .] the more prominent a particular [. . .] component [. . .] the greater the preponderance of the property characterized by that component. Notice too that |V| and |C| can characterise segments either alone or in combination.
>
> *(p. 151)*

'Prominence' of elements is expressed in terms of a head-dependent relation.

These dependency relations provide the tools to express a number of major segment classes in terms of combinations of |V| and |C|, as shown in (15):

(15)



Below the actual representations, we have indicated which classes of segments they represent. AE argue that the representations reflect a sonority ranking, going from left to right, in which the classes of voiceless fricatives and voiced stops are claimed to have equal sonority. Further distinctions (leading to separate representations for laterals, strident fricatives, etc.) will be discussed below. Note the use of complex structures that involve 'primary (or head) structure' like |V:C| entering into a dependency with another, 'secondary' structure, another instance of using the same element multiple times (within a gesture); see section 12.3.4.

In order to characterize the segment classes in (15) in a feature system of the SPE type (Chomsky and Halle 1968) we would need the features [voice], [consonantal], [continuant] and [sonorant], where DP uses just two single-valued features: the elements |C| and |V| and their interdependencies. However, pure reductionism was not AE's primary motivation for replacing major class and manner features by CV-complexes. They claim that their approach is more adequate than traditional binary theories in a number of respects. First, as seen above, by replacing binary features with structures of varying complexity, representations more adequately reflect the relative markedness of phonological major class and manner categories. In (15), the categories vowel and voiceless stop are the least complex, which reflects their relatively unmarked status. Fricatives are more complex than stops, and voiced obstruents are more complex than voiceless ones. This again reflects well-known and widely accepted claims regarding the relative markedness of these categories. Secondly, as also stated earlier, AE also claim that the array of structures provides an adequate characterization of the notion of relative sonority. Degrees of sonority correspond to the amount of 'V-ness' that a representation contains. (We could likewise define strength in terms of the amount of 'C-ness'.)

This is useful in the characterization of lenition processes (see section 5). Thirdly, AE claim that the structures composed of |C| and |V| provide a more adequate basis for the expression of phonological processes than traditional binary systems do. With reference to (15), AE note that these structures reflect an asymmetry in the behavior of 'voicedness', as opposed to 'unvoicedness'. If we assume (as most phonologists do) that phonological rules can only cause phonetic events by manipulating phonological units, the structures in (15) express that languages can spread 'voicing' but not the absence thereof. If this is empirically correct, representations as in (15) are superior to binary feature systems in which [+voice] and [−voice] have the same status.[22] Finally, the CV-constellations are constructed in such a way that affinities between the phonological categories that they represent are formally expressed. For example, in the structures in (15), an ungoverned |V| can be glossed as [(+)sonorant], whereas a governed |V| forms the equivalent of [(+)voice]. This particular example reveals that DP manages to express distinct but clearly related phonological categories in terms of a single primitive appearing in different structural positions, where traditional feature systems must stipulate a relation in the form of redundancy rules like [+sonorant] → [+voice]. In DP, [+sonorant] and [+voice] are manifestations of one and the same element, viz. |V|. The relation between these two categories is therefore inherent to the basic vocabulary.

Before we turn to a further discussion of the syntax of the categorical elements |C| and |V|, we will briefly discuss the other 'gestures' (element classes) in (13). First, we turn to the second subgesture of the categorial gesture, viz. the initiatory subgesture. DP advocates the idea that the traditional concept of phonation (involving glottal states and vocal fold vibration) is relevant to two different gestures. Vocal fold vibration (voicing) is, as we have seen, expressed within the 'phonatory' subgesture of the categorial gesture, whereas glottal state distinctions are incorporated in the initiatory gesture. This latter subgesture contains the 'glottal opening' element |O| ('aspiration') and two elements used for the description of different types of airstream mechanisms: |G| (for 'glottalicness', i.e. 'constricted glottis') and |K| (for 'velaric suction').

AE argue that the use of |O| is called for in three types of languages (AE: p. 188):

- Languages that have a voice distinction that involves more than two categories (e.g. Indonesian, which has voiceless, 'lax voice' and 'tense voice')
- Languages that do not seem to use voice but rather aspiration (e.g. Icelandic)
- Languages that have an opposition between voiced and voiceless sonorants (e.g. Burmese, which has this contrast for nasal and laterals)[23]

Proceeding with this sketch of DP, let us turn to the daughters of the locational subgesture. AE introduce the place elements in (16):

(16) DP place elements

|i| 'palatality, acuteness/sharpness'	|l| 'linguality'
|u| 'roundness, gravity/flatness'	|t| 'apicality'
|a| 'lowness, sonority'	|d| 'dentality'
|ə| 'centrality'	|r| 'retracted tongue root'
|α| 'Advanced Tongue Root (ATR)'	|L| 'laterality'

Not all these elements play an equally important role in the theory. The heart of the set of place elements is formed by the familiar 'aiu' subset, which plays a key role in the representation of vowels and consonants. Two further elements are added for vowels: centrality (already discussed above and perhaps redundant) and ATR (an element that we will return to in section 12.3.4). Here we will focus on the elements which are mainly or exclusively used for consonants (the right-hand column).

|l|, lingual, was motivated by Lass (1976) to capture the natural class of high front vowels and tongue blade and tongue body consonants, which he claims recurs in sixteen processes in the history of English.

|t| is meant to capture the contrast between apical and laminal coronals, while |d| distinguishes dentals from alveolars. Systems that have dentals and alveolars frequently distinguish these places also in terms of apical and laminal, although no system seems to have an apical/laminal distinction at either the dental or alveolar place of articulation. However, AE argue that in certain cases both |d| and |t| are necessary.

|r| is introduced to represent pharyngeal consonants. AE also consider using this element in vowels to capture the ATR/RTR distinction (AE: pp. 243–245). However, given the evidence that in many harmony systems the [ATR] value is dominant, AE suggest that another element, |α|, is needed for such systems.

|L| is introduced without too much motivation, simply to capture laterality, despite the fact that laterals are also captured in the phonatory gesture. One might say, however, that |L| is needed for lateralized segments such as lateral fricatives.

Here are some representative consonantal place representations:

(17)	{|u|}	{|l|}	{|l,i|}	{|l,u|}	{|l,u,a|}
	labials	dentals, alveolars	palatals	velars	uvulars

Note that the variety of elements that is used here in the representation of consonants somewhat weakens the idea that elements are used across the board, i.e. for both consonants and vowels (see the fourth assumption in (1) above). Both in DP and DP-inspired approaches (Smith 1988; van de Weijer 1996; Staun 1996b; among others) various proposals have been made to cut back the set of locational elements to the basic aiu-set. Also in Radical cv Phonology (van der Hulst in prep.) all the extra elements in (16) have been eliminated, with the resulting set being fully employed for both consonants and vowels.

The oro-nasal subgesture contains precisely one element, |n|, for 'nasality'. Recall that there also is a phonatory characterization of nasals {|V⇒C|}. This is comparable to the case of laterality, for which DP also proposes a phonatory representation (for laterals proper) as well as an element (for lateralization).

One might wonder whether DP really needs a nasality element, or, if it turns out that such an element is necessary, whether this element should occupy an entire subgesture by itself, which seems to have been proposed on the basis of general considerations of symmetry. With respect to the first question, AE argue that nasal consonants not only form a natural class with other sonorant consonants by sharing certain characteristics in their categorial (particularly phonatory) representations, but they also form a natural class with nasalized segments, which may have different specifications in the categorial gesture. In order for this latter natural class to be reflected by the DP representations of the segments in question, AE argue that we need a separate component, |n|.

Before we return to the 'phonatory' (i.e. major class/manner) subgesture, let us briefly look at AE's proposals for the tonological gesture. In their excursus on representations for tonal distinctions, AE make the intriguing suggestion that the elements |i| and |u| (as part of the tonological gesture) could be employed for high and low tone, respectively.

> [W]e propose that the appropriate representations for the two tonal components are [. . .] |i| and |u|. In other words, we are suggesting that |i| and |u| in the tonological gesture bear the same relation to |i| and |u| in the articulatory gesture as |V| in the categorial gesture does to |a| in the articulatory gesture. [. . .] That is, |i| involves (relatively) 'high frequency' and |u| (relatively) 'low frequency'; whether this is interpreted as high (or low) F_0 or as concentration of energy in the higher (or lower) regions of the spectrum depends on the context – i.e. gesture – in which it occurs.
>
> *(p. 273)*

What is most noticeable in this proposal is the idea to use the same elements, viz. |i| and |u| in two different gestures. To emphasize that this strategy is present in the AE proposals, we will here also quote AE on their suggestion concerning the identity of |a| and |V|.

> [T]here is clearly a relationship between |a|, as a component within the articulatory gesture, and |V|, as a component of the categorial gesture. Consider the acoustic glosses which we have given the two components: |V| corresponds with maximal periodicity, and |a| with maximal sonority. Vowels, by virtue of their periodicity are the most sonorous of the categorial segment-types, while open vowels are the most sonorous within the class of vowels. [. . .] The open unrounded vowel, then, might have {|V|} both as the representation of the categorial gesture and of the articulatory gesture.
>
> *(p. 215)*

The importance of these quotes is to show that AE suggest the strategy to employ the same elements *in different (sub)gestures* (which needs to be distinguished from using the same element more than once within a gesture), thus deriving similarities in phonetic interpretation while attributing the differences to the fact that the '(sub)gestural location' of an element has a bearing on the phonetic interpretation as well.

This shows that DP offers two possibilities for reducing the number of primes. Firstly, fewer primes are needed due to the dependency relation. Two traditional features can be replaced by the dependent and head occurrence of a single prime, e.g. |V| for [voice] and [sonorant]. Secondly, fewer primes are needed given grouping. One particular element may occur in various groups, each time with a different phonetic interpretation and thus replace two or more features.

In section 12.3.4 we will elaborate on this reduction strategy, which forms the foundation of Radical cv Phonology.

12.2.2.4 Minimal specification and polysystematicity

Even though the adoption of unary features pre-empts the notion of underspecification in many ways, it does not become inapplicable. Anderson advocates a strong minimalist view with respect to the specification of phonological information, which must be strictly contrastive. All redundant, predictable properties should be eliminated from the representation. Underspecification becomes relevant when we consider positional phonotactic restrictions, as for example in the well-known case of English initial clusters. In a trisegmental cluster like /spr/ the initial segment, if consonantal, can only be /s/, which means that all properties of this segment, except its consonantality, are predictable. Likewise, the second segment (a voiceless stop) and third segment (an approximant) have many predictable properties. Without spelling out what the minimal representation in terms of components would be, it seems clear that very few elements are required.

It is important to note that (in general, not just in DP) the use of underspecification undermines the traditional notion of the phoneme as a unit that generalizes over allophones that occur in different positions, being in complementary distribution. Such a rejection is masked by the use of terms like 'archiphoneme'. Rather, it leads to a type of analysis in which each position in the string of segments has its own contrastive set of oppositions (its own segment system, so to speak). This means that phonology is *polysystematic* (as recognized in the Firthian approach; Firth 1948). For example, if a language limits syllable-final consonants to plain voiceless stops, the relevant position only allows a contrast between whatever the plain voiceless stops are that the language allows in terms of place. If this is labial, coronal or dorsal, then a final 'k' can simply be represented as {consonantal, dorsal}. However, an initial 'k' might contrast with all other consonants and might therefore have a richer representation, e.g. {consonantal, voiceless, stop, dorsal}. The polysystematic view holds that these two sets of features are independent and *not* unified under a joined concept of 'the phoneme /k/'. Nevertheless, these two sets are mapped onto phonetic events which happen to be very similar. The classical notion of the phoneme formally expresses this phonetic similarity which, as argued by Pike (1947), provides a natural basis for an economical alphabetic writing system. However, Anderson sees this traditional notion of phoneme as not being a genuine phonological entity.

In conclusion, segments in all positions of the syllable have their own sets of oppositions. Segments in a given position are specified minimally to distinguish them from other segments that can occur in the same paradigmatic slot. Furthermore, in any such system one member can always be specified as the null option (i.e. without any elements).

Anderson extends the use of underspecification to linear order. We return to this point in section 12.4 where we discuss the DP approach to syllable structure.

12.3 Developments in DP

In this section and the next we discuss several developments that have taken place in DP, especially in the characterization of segmental structure. We will organize these developments according to the (sub)gestures they apply to.

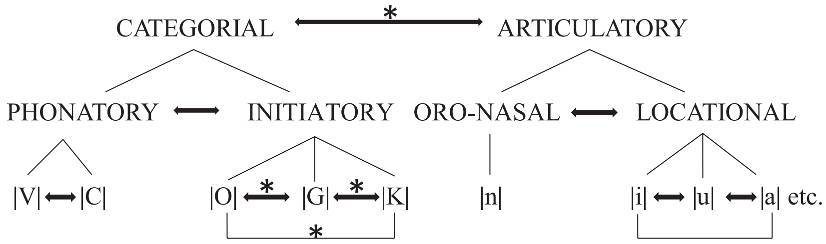
12.3.1 Developments with respect to inter- and intrasubgestural dependency

Standard DP used the possibility of allowing subgestures to enter into dependency relations, but this was not fully exploited. This is schematically summarized in (18), where an asterisk indicates that no dependency relations are proposed between the units connected by the bidirectional arrow.

In (18) we also encode that there are no dependency relationships between the two main higher gestures: there are no circumstances under which segment types are distinguished by means of a difference in the dependency relation between the components of the categorial and articulatory gestures.

It is unclear why AE use precisely the dependencies illustrated in (18) and no others. In an attempt to restrict the DP model, Davenport and Staun (1986) argued to dispense with inter-subgesture dependency. They show that once the glottal opening component |O| is assigned to the major class/manner ('phonatory') subgesture and a new component |i| ('initiator velocity', expressing the direction of airflow) is assigned to the initiatory subgesture, there no longer is a need for dependency relations between the phonatory and the initiatory subgestures. We refer to Davenport and Staun's (1986) work for further discussion of this point, and the ramifications of their proposal for the DP framework.[24]

(18)



12.3.2 Developments with respect to the oro-nasal subgesture

Noting that DP expresses nasality in two ways (see section 2.2), Davenport (1995) proposes to dispense with the component |n| altogether. This implies that the categorial characterization of nasality 'survives', although Davenport's proposal is that nasality is not expressed in the major class/manner ('phonatory') subgesture (i.e. not in terms of a specific |C|/|V| combination), but as a separate component |N| in the initiatory subgesture. So, in a sense, Davenport's proposal is a compromise between the two 'old' ways of expressing nasality in DP. We refer to Davenport's article, which shows that the dual representation of nasality leads to unsatisfactory results in DP.

12.3.3 Developments with respect to the initiatory subgesture

Davenport and Staun (1986) maintain an initiatory subgesture, which contains components for airstream distinctions: |I| 'egressive airflow' (not present in AE), |G| 'glottalicness' and |K| 'velaric suction'; |O|, which forms part of this subgesture in AE, has been moved to the phonatory subgesture in their model. However, their proposal has not been worked out in further detail, as far as we know, and so it remains 'food' for further thought on the issue of intrasegmental structure within DP.

It is noteworthy that research in DP has not developed a separate 'laryngeal' gesture that would capture voicing, aspiration and glottalic constriction (as in most Feature Geometry models). It is also noteworthy that Feature Geometry proposals have generally not proposed a class node with features for initiation, i.e. for ingressive sounds like implosives, and clicks or egressive sounds like ejectives. Segments of the latter type are usually expressed with laryngeal features or as complex segments with a double articulation (see Sagey 1986).

12.3.4 Developments with respect to the Major class/Manner 'phonatory' subgesture

We will now turn to a more extensive evaluation of the organization of the phonatory subgesture and argue that the 'syntax' of CV combinations is not clearly defined in AE's version of DP, a point also emphasized in den Dikken and van der Hulst (1988), who offer an alternative which can be seen as an important step in the development of Radical cv Phonology (van der Hulst 1994, 1995, 2005, in prep.).

For convenience, in (19) we repeat the set of distinctions built from |C| and |V| which AE propose as a kind of core set:

(19)

	{|V:C|} vcl. fric					
{|C|} vcl. stop		{|V:C⇒V|} voi fric	{|V⇒C|} nasal	{|V⇒V:C|} liquid	{|V|} vowel	
	{|C⇒V|} voi stop					

The core of this set is formed by the five different basic structures that are composed of two elements:

(20)	{|C|}	{|C⇒V|}	{|V:C|}	{|V⇒C|}	{|V|}
	stop	voi stop	fricative	nasal	vowel

As we see in (19), this set can be expanded by adding a secondary instance of a basic structure in dependent position.

From the viewpoint of generative power, one would like to know exactly what the set of possible C/V combinations is that includes primary and secondary structures. AE do not address this issue explicitly. Rather, as seems motivated by the attestation of potential manner contrasts, they continue to add new structures, more or less in an ad hoc way (even though they provide cogent arguments for each individual structure that they propose). For example, AE add the following more complex representations to capture further distinctions:[25]

(21)	{|V:C⇔V|}	{|V⇒V:C⇒C|}	{|V:C⇔V⇒C|}	{|V:C⇒C|}
	fricative trill	lateral	voiced lateral fricative	non-sibilant fricative

Here we even see the use of three levels of structure for the two categories in the middle.

The argumentation that AE provide in favor of these representations is based on attested natural classes. Fricative trills may pattern with voiced fricatives in conditioning phonological processes (AE give 'Aitken's Law' as an example). Given the representations in (21), the relevant natural class can be represented as in (22):

(22) {V:C⇒V}

Lateral liquids, of course, must be distinguished from r-sounds, which motivates the second structure in (21). AE write:

> [L]aterals are phonetically unique, as far as the phonatory sub-gesture is concerned, in having effectively two manners of articulation. While there is a stricture of open approximation at one or both sides of the mouth (at least for sonorant laterals), there is also closure in the centre of the oral tract. [. . .] Essentially, then, the |C| node characterizes a secondary [. . .] stricture type within the phonatory sub-gesture.
>
> *(p. 163)*

The dependent |C| in laterals expresses the fact that laterals may pattern with stops. In traditional feature systems, there is no direct way to express such a class without introducing the feature [continuant] in laterals, which is redundant since laterals are already uniquely characterized as [+lateral].

The extra dependent |C| in the third representation in (21), then, also adds laterality to the fricatives (p. 164). The fourth structure reflects the distinction between sibilant and non-sibilant fricatives:

> [. . .] /s/ may be interpreted as the optimal fricative phonetically; acoustically it shows the 'simplest' combination of consonantal and vocalic properties, while the other fricatives involve energy reduction in various frequency bands. In comparison with the sibilants, then, the other fricatives display extra /C/-ness.
>
> *(p. 166)*

Even though AE carefully motivate the structures in (20) and (21), formally capturing many relations between different sound classes that must be stipulated in traditional feature theories, questions can be raised concerning the restrictiveness of their approach. The 'syntax' underlying combinations of components (|C| and |V| in this case) is not explicitly defined, i.e. we do not know what the total set of possible dependency structures is. Clearly, AE assume that the syntax is, in a sense, recursive, so that structures that have been formed can be input to further combinatorial structures. However, given that this recursive syntax allows, in principle, many other structures, we must conclude that AE make no serious attempt to come to grips with the notion 'possible phonological segment'. Arguably, the notion of possible segment does not play a decisive role for AE. Their approach allows one to conceive of structures of various degrees of complexity and the only relevant concern would then be to predict that more complex structures imply structures of lower degrees of complexity within a given language (within a given position).

While this is a valid position, den Dikken and van der Hulst (1988) nonetheless make a proposal with respect to the use of the components |C| and |V| that imposes a general limitation on the complexity of CV-structures. The initial idea in this proposal (based on van der Hulst 1988, discussed in section 2.2.2) is that each component can occur *at most* twice. In several articles and in work in progress, van der Hulst has developed this initial proposal, trying to maintain a systematic and 'controlled' set of structures in which each structure is actually used to express attested contrasts; this is the theory of Radical cv Phonology (RcvP) (cf. van der Hulst 1995, 1996, 2000, 2005, 2015a, 2015b, in prep.). Recall that AE explored the use of the same elements in different subgesture (see section 3.1 above). In RcvP, this idea is pushed to its logical extreme. In addition, the proposal is that there are only *two* elements. Somewhat arbitrarily, RcvP adopts the labels |C| and |V| for these two elements. In each gesture, these two components allow a four-way distinction in phonological classes: C, C;V, V;C, V.[26,27]

These structures will receive different interpretations depending on the syllabic positions that they occur in:

(23)	Onset head	C	C;V	V;C	V
		stop	stop strident	fric.	strident fric.[28]
	Onset dep.	nasal	liquid	rhotic	glide
	Rhyme head	high	high mid	low mid	low
	Rhyme dep.	nasal	liquid	rhotic	glide

The syllabic structure also has a four-way distinction (C, C;V, V;C, V), encoded in its basic template, which maximally allows a branching Onset and a branching rhyme.

While this proposal allows a reduction to four basic structures, there does seem to be a need for some further finer distinctions which thus call for secondary occurrences of the C and V components. We have seen that the use of secondary structures was already present in (21). Although Anderson (2011c) does not present a complete outline of the DP elements and their structures in phonological segments, he *explicitly* recognizes a distinction between primary and secondary occurrences of elements, which represents a major innovation compared to AE. Revising the combinatorial system in (18) and (19), he proposes to represent nasality and voicing in terms of secondary occurrence of the C and V elements:

(24)	{V;C {c}}	{V;C}	{V;C{v}}
	nasal	lateral	rhotic

Anderson (2011c: 114)

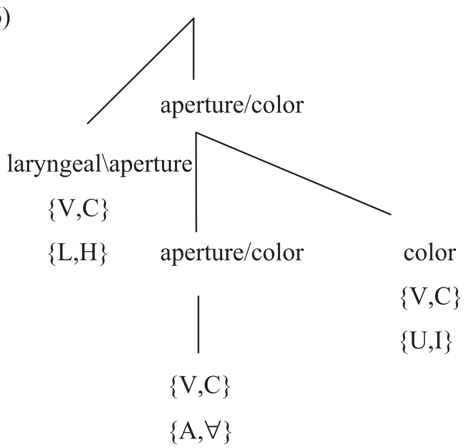
(25)	{C;V{v}}	voiced fricative
	{C{v}}	voiced stop

Anderson (2011c: 362)

The idea to use elements in a secondary role (which is also a trait of RcvP) deserve further exploration.

We conclude this section with one example from van der Hulst (in prep.). The RcvP model, as we have discussed, postulates two antagonistic elements in each class. This raises the question how the triangular approach to location in the AIU approach is incorporated in this model. Van der Hulst suggests that the traditional view, which regards |u| and |i| as 'colors' and |a| as 'sonority', suggests that |u| and |i| belong to one class (which we may call 'color' for convenience), whereas |a| belong to another class (which we can call 'sonority' or 'aperture'). However, this implies that |a| must have an antagonistic counterpart, which van der Hulst represents as |∀|. The element labels used here are for convenience only, because the real elements are |C| and |V| in both classes:

(26)

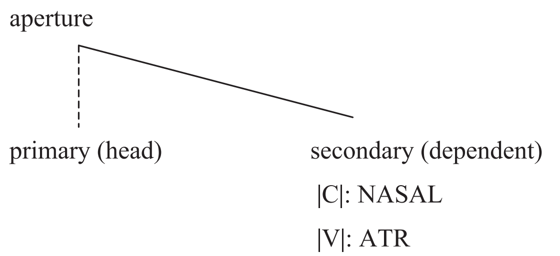


The labels 'aperture' and 'color' as well as 'laryngeal' are merely mnemonic shorthands for structural representations that indicate that the 'aperture' node is the head, taking 'color' as a dependent complement (indicated by the '/') and 'laryngeal' as a dependent adjunct (indicated by '\') (see Anderson 2011c: 355, and below for the use of this 'slash' notation). Likewise, the use of the traditional element labels (A, I, U, L, H) simply serves the purpose of reminding the reader how the C and V elements in the different classes are phonetically interpreted.

By recognizing a fourth basic element, namely |C| in aperture, the RcvP model converges on six elements, just like certain recent versions of Government Phonology (see Scheer and Kula, this volume: Chapter 9, section 2.1), where the so-called |ʔ| element correlates with the |C| element in RcvP's aperture.[29]

Van der Hulst then proposes that both elements can occur as a secondary (dependent) element, which is a dependent to the aperture unit in (26):[30]

(27)



For vowels (or nuclei), the secondary |V| is interpreted as pharyngeal (ATR), whereas one proposed interpretation of secondary |C| is NASAL, which would imply that the secondary elements denote the two non-oral cavities, pharyngeal and nasal, respectively.[31,32]

The RcvP model explores the use of secondary elements for all three element classes.

12.4 Suprasegmental structure

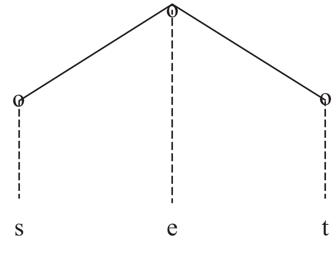
In DP it is assumed that the syllable is the basic unit for expressing phonotactic restrictions, and that, in addition, several phonological processes also motivate the syllable as a domain. Syllables are headed constructions, because they are "characterized by the presence of an atomic element, the syllable peak, in whose absence there is no syllable" (Anderson and Durand 1988: 9).

A simple syllable such as /set/ can be characterized by the following two statements:

(28) a. government relations: s ← e → t (e governs s and t)
b. precedence relations: s < e < t

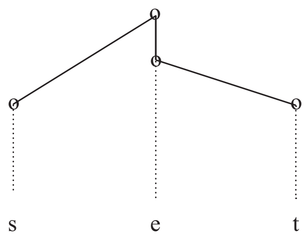
In a dependency graph, all segments are represented as nodes, which are connected by lines. Head nodes are represented higher on the vertical axis:

(29)



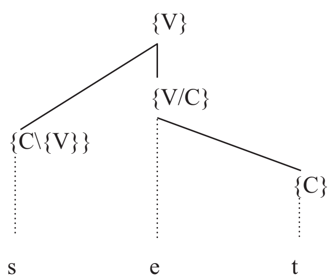
In this structure the ONSET and coda consonants are equal dependents of the nuclear head vowel. Anderson suggests that the following structure, which introduces subjunction, is also consistent with the basic principles of dependency grammar:

(30)



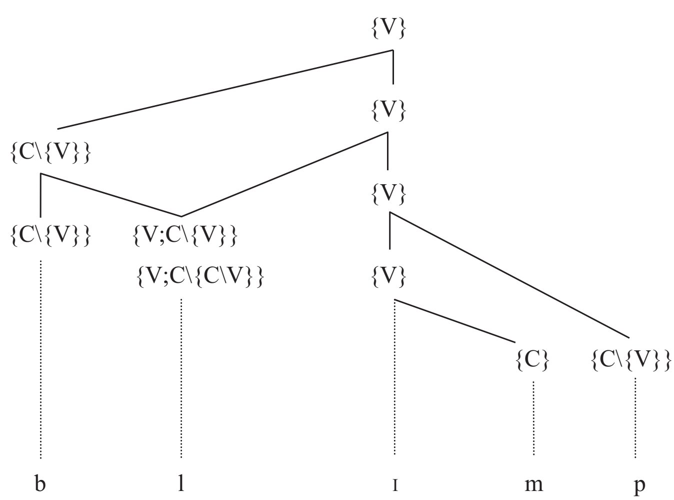
Here /e/ is dominated by two nodes, one subjoined to the other. The /t/ is taken to be a complement that is selected by the lax vowel, which requires a following consonant. The /s/, on the other hand, is an adjunct. Anderson (2011c: 83ff.) introduces the following notation to represent the various nodes:

(31)



The '/' indicates 'looking for a complement', while the '\' notation stands for being an adjunct to what is to the right of '\'. Anderson's approach only uses binary structures, so for more complex syllable types additional structure is needed. This is illustrated in (32):

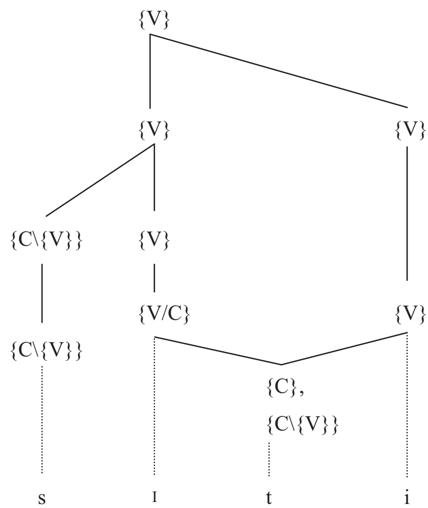
(32)



The second consonant /l/ ({V;C}) is, at the same time, an adjunct to the /b/ ({V;C\{V}}) and to the following vowel /ɪ/ {V;C\{C\V}}). Likewise, the final consonant /p/ is an adjunct to the 'rhyme' that is formed by the vowel and following consonant.

Dependency graphs also permit one daughter to be dependent on two heads, which creates a structure that appears to correspond to the notion of ambisyllabicity (cf. Kahn 1980; among others):

(33)



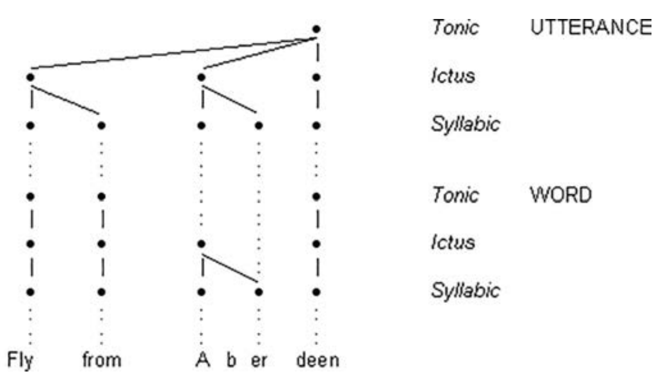
Here we have also included the dependency relation that represents 'foot structure'. Indeed, early work in DP anticipated the essence of metrical theory by representing 'stress' as an exponent of a dependency relation between two syllables.

Anderson (1986b) proposes an interesting constraint on syllable representations:

(34) The Dependency Preservation Condition
Dependency relations are preserved, where possible, throughout a derivation (and in diachronic changes)

Anderson introduces this condition as part of his syllabification algorithm, to ensure that dependencies introduced by earlier rules are not undone or reversed by later rules. We note that this principle anticipates the Projection Principle proposed in GP (see Kaye et al. 1990).

Within the expression plane, Anderson makes an intraplanar distinction between word structure and utterance structure, which is more or less equivalent to the distinction between lexical and post-lexical structure. Here we reproduce a diagram from Anderson (1986b) which illustrates this distinction (and which abstracts away from many details of node labeling):



Each word has its own dependency structure, capturing syllabic structure and stress. Then words are gathered into an utterance structure, which in particular cases imposes a post-lexical foot-like structure that is reminiscent of the so-called Abercrombian foot in grouping syllables that belong to different words.[33]

12.5 Rules in DP

This section discusses how phonological alternations are represented in DP. Proponents of DP do not always agree on which rules should be accounted for in the phonology, and which are merely lexical idiosyncrasies. Recently, Anderson (2014) addressed this topic, making the claim that there are no phonological rules, except structure-building redundancies. There are alternations manifested in pairs of morphologically related lexical items, and there are adjustments when morphological units are put together, expressed in the interface between morphology and (lexical) phonology – i.e. the morphophonology. Similarly, there are adjustments at the lexical–utterance interface. There are no phonological mutations or shifts, except diachronically. Nor do proponents of DP always agree on the role of abstractness, i.e. the specific question to what extent underlying representations should be allowed to diverge from the phonetic surface. However, the fact that phonology is substance-based militates against ‘ghost segments’ that never reach the surface as well as empty syllabic positions.

Rules in DP are generally quite comparable to normal autosegmental spreading operations, with the obvious proviso that only elements (corresponding to the ‘marked’ binary feature values) can be spread (or be referred to in constraints). Where effects arise that do seem to require such rules, additional machinery (e.g. in the form of constraints) is necessary. Similarly, the elements posited in DP can be used in constraint-based frameworks (Prince and Smolensky 1993 [2004]) without difficulties. Here the question arises of what the set of elementary features is, but this is fundamentally a different question of whether the head-dependency relation can be used among such features.

DP assumptions about segmental as well as suprasegmental structure are helpful in an understanding of processes of vowel harmony. With respect to segmental features, DP makes strong predictions about what types of harmony are found (viz., ones that are based on existing elements) and what types are not found (viz., ones that are based on the ‘negative values’ of unary elements). It also helps to characterize the targets of harmony as syllable heads, while consonants play a secondary role (see e.g. van der Hulst and van de Weijer 1991, 1995; van der Hulst to appear).

Another process that is particularly elegantly captured in DP is neutralization (see section 12.2.2.2 above, and e.g. Anderson and Ewen 1981; Staun 1985). For vowel neutralization, we can think of different sets of vowels appearing in different positions: stressed vs. unstressed, oral vs. nasal, in roots vs. in affixes, where typically the vowels in the latter conditions form a subset of the vowels in the former condition. Although languages differ in their patterns, in all cases that we know of, the reduced set can be described as lacking an element and/or the head-dependency relation that is present in the fuller set. It is also important to note, again, that the analyses can be conceived of as rules or as the result of constraint interaction. For consonant neutralization, final devoicing was mentioned above, which favors an unmarked consonant type over a marked one.

Thus, in many cases, the relatively constrained tool set of DP results in more elegant accounts of segmental processes. Such is also the case with processes like diphthongization (e.g. /eː/ → /ei/ → /ai/), merger of vowels (/ai/ → [e]), vowel lowering and raising rules, and breaking (see e.g. Anderson and Ewen 1987; Anderson 1986a; Colman 1987, 2005; Lass 1987; Rennison 1986, 1987, 2014).

DP offers the extra mechanism of rules (and constraints) based on the dependency relation alone, i.e. affecting the headship of one of the elements that enters into a dependency relation with another element. Rules of this type elegantly account for processes that are more difficult to describe using traditional distinctive features, e.g. vowel raising and lowering (chain shifts), or neutralization of vowel contrasts in particular positions.

Finally, lenition (either as historical process or in synchronic phonology, see e.g. Gurevich 2011) is hard to capture in frameworks based on binary features, since a number of different features ([voice], [consonantal], [continuant], [sonorant], etc.) are involved in what appears to be a unified phenomenon. Representations like those in (19) above are eminently suited to capture lenition as a shift in the preponderance of the element |V| (see e.g. Ó Dochartaigh 1979, 1980 for an analysis of Celtic lenition in the DP framework).

To express certain types of chain processes, DP allows a mechanism called resolution, which was already proposed in Anderson (1973):

(35)	a.	Add	B to A =	AB	= *A*B
		Add	B to *A*B =	*A*BB	= A*B*
		Add	B to A*B* =	A*B*B (or ABB) = B	

b. A > *A*B > A*B* > B
(each step: Add B)

This schema applies as follows:

(36)	Add V to C	⇒	C;V	(high vowel becomes high mid)
	Add V to C,V	⇒	V;C	(high mid vowel becomes low mid)
	Add V to V;C	⇒	V	(low mid vowel becomes low)

This schema allows the representation of processes that involve the apparent deletion of elements. It can be applied both to the vowel-related shifts (e.g. the Great Vowel Shift in the history of English), or to consonant-related phenomena such as lenition.

12.6 Related approaches

In van der Hulst and Smith (1982), the ideas of DP were presented in the context of an overview of recent non-linear developments in Generative Phonology. Although these ideas have remained largely unnoticed, three major hallmarks of DP (monovalency, grouping and intrasegmental dependency) have all, in various degrees, been incorporated in various other approaches, including 'mainstream' Generative Phonology, especially in the development of Feature Geometry, a movement that started around the early to mid-1980s. Here we mention the crucial parallels.

Feature theories in mainstream Generative Phonology have also appealed to unary features, but in a weaker form, by proposing that only *some* features are single-valued. For example, various scholars have suggested that [round] is single-valued (e.g. Steriade 1987). Itô and Mester (1986) argued that [voice] is a single-valued feature. Goldsmith (1985, 1987) went even further and proposed a system in which both [round] and [low] are single-valued, with the proviso that the scope of [low] is extended to low and mid vowels. In his system, [back] is still binary. The strong version of this claim says that *all* features are single-valued.

This strong position was precisely what Anderson and Jones proposed. The use of unary elements, more specifically the use of the triangular IAU set, was adopted in the approach of Schane (1984, et seq.) who applied these elements to vowel processes, in particular monophthongization and diphthongization. Schane did not employ dependency, but instead used an additive mechanism. Van Nice (1991) proposed an extension of PP in which the elements |i| and |u| are grouped under a single node. Similar proposals were made in Ewen and van der Hulst (1985, 1988) and van der Hulst (1989) within the context of DP. Further applications of PP can be found in Hayes (1989) and Broadbent (1999). The latter adds a dependency relation and thus removes the idea of stacking particles, turning this essentially into a variant of DP.[34] The use of so-called empty nuclei, a hallmark of Government Phonology (GP; Kaye et al. 1985, 1990), is not acknowledged in DP, which, given its substance-based approach, cannot make reference to units that have no substantive correlate.

The pivotal aiu-set of elements was also adopted in GP, which in addition also introduced the use of dependency relations between elements. Both DP and GP emphasized the perceptual nature of the elements, as well as the idea that the elements generalize across vowels and consonants. That said, both models went through a similar development of proposing additional elements, sometimes elements that would only be used for the representation of consonants. GP has reverted to a simpler set of six elements, while one variant of DP, namely RcvP, makes a very similar proposal (see section 12.3.4 above). A point of potential difference between GP and DP could be that the former insists that each element can be independently phonetically realized. In spirit, this demand would seem to square with the substance-based approach of DP, but the independent realization has simply not been taken to be a condition on elementhood in DP; nor is it clear to us why such a condition would have to be imposed. GP claims to be a theory about the computational system that underlies phonology and as such it is stressed that phonetic factors can play no role in establishing a phonological model. DP does not make such a claim and, in fact, by making grounding a cornerstone of the entire enterprise, it could never be impervious to phonetics. But in fact, GP's basic elements are firmly rooted in acoustics, just as in DP. In practice, GP and DP come to a very similar conclusion about what phonology is about, with the exception of DP's denial of empty nuclei, which drives a wedge between both models that is caused by GP's non-commitment to phonetic substance in all respects.

For a close comparison of DP and GP versions of element theory, we refer to den Dikken and van der Hulst (1988) and van der Hulst (2016a). In recent years GP has developed a use of headedness which is perhaps different, in that elements are used as either headed or non-headed, irrespective of whether or not they occur in combination with other elements. This introduces a kind of diacritic headedness which we do not find in DP (see Scheer and Kula, this volume: Chapter 9). DP and GP also converge on the rejection of constituent structure in favor of a strictly relational approach in terms of head-dependency relations.

Certain proposals in GP have also developed the idea of an intrasegmental grouping. We refer to Chapter 9 in this handbook for a discussion of various proposals. One such proposal, developed in Kula (2002), while placed within the context of GP, proposes an element 'geometry' that incorporated various aspect of standard DP proposals as well as of RcvP.

There is furthermore a striking parallel between DP and GP regarding the rejection of constituency. For example, with respect to syllable structure DP does not appeal to constituents such as onsets and rhyme, or even a constituent syllable. Dependency graphs do not represent constituency. We here draw attention to the fact that a similar stance is taken in current versions of GP (see Scheer and Cyran, this volume: Chapter 10) which claim to abandon constituency in favor of so-called lateral relations. It seems to us that the representation of 'syllable structure' and other relations in terms of these lateral relations between segments as basic units comes close to being a variant of the dependency approach.

We conclude that, viewed from a certain distance, DP and GP have come very close, although there are still differences that may be difficult to bridge, such as the rejection by DP of empty elements or the matter of whether or not elements should be independently pronounceable.

With reference to Feature Geometry proposals, we observe three parallels.

The idea that one set of elements can generalize over consonants and vowels (while not fully adhered to in the original proposals in Anderson and Ewen 1987, but restored in later DP work by others) also occurs in Feature Geometry models; see Hume (1994); Clements and Hume (1995); Padgett (2011); among others. This idea was also present in the earliest work on binary features (Jakobson et al. 1952), but had been abandoned in Chomsky and Halle (1968). A return to using the same features for consonant and vowel distinctions can also be seen in proposals to combine one set of features for tone and phonatory categories (cf. Yip 1980; Duanmu 1990; and Bao (1990), following the spirit of Halle and Stevens 1971).

As discussed in section 12.3.2, DP proposed a dual representation for nasality, i.e. in terms of a C/V combination and in terms of a separate element for nasality. Proposals within Feature Geometry have sometimes also adopted a separate node for the feature nasal (cf. Sagey 1986, 1988). Piggott (1990, 1992) proposes a 'velic class node' dominating only [nasal]. In addition, he adopts a node 'spontaneous voicing', which may also dominate a feature nasal. The duplication of nasality in Piggott's model bears a clear resemblance to the way DP treats nasality, but its precise status remains a topic of controversy.

Finally, with reference to Feature Geometry, it is obvious that the DP notion of gesture is completely parallel to the class nodes that were introduced in the work of Clements (1985) and Sagey (1986).[35]

12.7 Conclusion

In this chapter we have reviewed the initial proposals and later developments of DP. We have highlighted the following properties of this approach:

- The use of unary primes (DP, shared with GP, PP)
- The use of dependency relations between primes (DP, shared with GP)
- The use of grouping (DP, shared with Feature Geometry and some versions of GP)
- The occurrence of elements in more than one group
- The replacement of constituency by head-dependency relations (shared with GP)
- Polysystematicity, i.e. a rejection of the phoneme as an abstract unit that generalizes over phones that are in complementary distribution
- Strict minimality: representations are stripped of all redundant properties, including linear order (within syllables), where this order can be derived from dependency relations and general principles of linearization (mainly based on sonority)

The use of the same elements in different gestures, for which the seeds were planted in Anderson and Ewen (1987) was pushed to the extreme in Radical cv Phonology (which otherwise embraces all the traits of DP),[36] which uses its recognition of grouping to reduce the set of elements to just two.[37] Since these two elements occur in three gestures, a six-way division results, which parallels recent proposals in GP.

Anderson and Ewen (1987), based on nearly two decades of previous work, present a complete research program for phonology which anticipated some of the major developments that took place in the field of phonology at large. The approach puts emphasis on the explanatory strength of a restricted representational system and on grounding phonology in phonetics. The least developed aspect of DP is its rule component, because a derivational account of alternations is not taken to be part of the synchronic phonology; synchronically there is simply a morphological alternation. In spirit, DP favors a surface-oriented approach, avoiding abstract (non-substance-based underlying or lexical representations) and (extrinsic) rule ordering.

Notes

1 Both these sections recapitulate, with modifications, parts of den Dikken and van der Hulst (1988).
2 Full disclosure: the authors of this chapter, which focuses on the work of John Anderson as the originator of Dependency Phonology (in the context of his adoption of Dependency Grammar for all modules of the grammar), subscribe to the basic tenets of the Dependency approach.
3 Anderson places morphology in the lexicon. In this component the units are combinations of basic phonological and basic syntactic units; see Anderson (2011a, 2011c).
4 Differences between the planes can also be due to how the primitive elements combine, as well as to how these planes interface. With respect to the former point, we observe that while recursion is possible in both syntax and phonology (see van der Hulst 2010), it is much more widespread in syntax.
5 We may speculate about the question whether the head-dependency relation is a purely linguistic characteristic, or that it belongs to a more general cognitive domain. Humans surely possess strong systems of perception and association, which helps them to make sense of the world, which typically involves many parts and in which relations between parts are important. From birth onwards, infants will learn that in any environment some parts are vital, and some merely 'background noise'. They quickly learn (or perhaps know innately) that some parts are worth focusing attention on, and some parts may be discarded.
6 Progress in segmental phonological theory in general has been halting, we believe, as a result of the rise of Optimality Theory. We hope that renewed interest in this field, e.g. based on advances in cognitive science, will pay special attention to the dependency relation.
7 A prepublication appeared in 1972 in *Edinburgh Working Papers in Linguistics*. This paper did not propose the second principle in (1), which was introduced later, following Lass and Anderson (1975).
8 See van der Hulst (2016a) for an overview of the unary/binary 'debate', and van der Hulst (2016b) for references to some earlier proposals for unary features.
9 See Sanders' (1972) simplex feature hypothesis; and see van der Hulst (2013) for some earlier precedents.
10 Whether voiceless for obstruents is unmarked in all positions could be a matter for debate, given the tendency for intervocalic voicing.
11 Steriade's contrastive specification theory would only leave non-contrastive values unspecified (Steriade 1987).
12 Various notations have been used for unary features, such as bold lowercase. We will use lowercase. Elements are enclosed in vertical lines.
13 The following is partly based on den Dikken and van der Hulst (1988).
14 Their choice of three units resembles the adoption of two 'colors' and 'sonority' in Natural Phonology (see Donegan 1978, which in turn echoes Jakobson's 1968 color and sonority axes). The triangular idea also resembles Stevens' quantal distinction as well as the proposals in Wood (1975, 1979). See van der Hulst (2015b).
15 Different notational systems have been employed by different authors, both for single elements and for combinations of elements. Here we use curly brackets.
16 Of course, many other notations can be used. In GP, for example, the head element is underlined.
17 Government Phonology makes the same claim (see Kaye et al. 1985), but they derive it from the internal logic of their theory.

18 Van der Hulst also proposed that for each element we expect that its head occurrence automatically entails the dependent occurrence, unless the presence or absence of the dependent is contrastive in a system.
19 We will return to the idea that ATR is a manifestation of the |I| element in dependent position later on.
20 In treatments of vowel harmony in Turkish the back unrounded vowel, which harmonizes for both roundness and frontness, would for that reason alone be specified as 'empty'; see van der Hulst and van de Weijer (1991).
21 This point is also acknowledged in Anderson (2011a, 2011c, 2014).
22 In this particular case, voicing, there is a significant literature claiming that the phonology needs reference to both values of voicing. See, for example, Wetzels and Mascaró (2001) and Chapter 15 of this book.
23 This has also been suggested in other work, such as Lombardi (1991), and it is supported by the fact that languages that have a voicing contrast for sonorants invariably also have an aspiration contrast for stops, as well as by the fact that in English approximants are partially devoiced in clusters of voiceless stops followed by an approximant: aspiration in vowels (*key*, *tin*) is phonetically similar to devoicing in approximants (*clean*, *twin*).
24 AE also exploit the possibility of allowing variable dependency between the two subgestures of the articulatory gesture. Arguably, one could be skeptical about the two distinctive degrees of nasalization, however.
25 In the second and third case AE do not indicate whether the (mutual) dependency relations are hierarchically ordered.
26 Van der Hulst (2005) rejects mutual dependency.
27 As noted in Anderson (2011c), replacing all elements by |C| and |V| in all classes is an instance of plane-internal structural analogy. Anderson resists the idea that all classes need to make the exact same set of structural distinctions, as van der Hulst seems to imply. This is comparable to his rejection of arguing that all phrase types in syntax must have the same structure, as originally proposed in X-bar syntax.
28 The representation of the strident/non-strident distinction for fricatives remains a problem. Also, contra Anderson, van der Hulst does not represent voicing in the manner class: he expresses this with a secondary v-element, but still in the categorical gesture (see the following discussion).
29 This convergence is discussed in more detail in van der Hulst (2016a). In a sense this fourth element, with reference to vowels, also restores the cold vowel that was proposed in Kaye et al. (1985).
30 Van der Hulst (in prep.) also discusses secondary occurrences of the elements in the other two gestures. All gestures generalize over consonants and vowels. The laryngeal gesture represents phonation distinctions for consonants and tonal distinctions for vowels. He proposes that, in its representation of tonal distinctions, this gesture can adjoin separately to the whole segmental structure in order to encode the 'autosegmental' nature of tonal properties.
31 Another phonetic interpretation of 'pharyngeal' that can be considered is RTR. The interpretation of pharyngeal as either ATR or RTR is taken to be 'areal'. By subsuming both phonetic interpretation under one element, it is explained why no language uses both contrastively; see van der Hulst (to appear).
32 For consonants, in line with Anderson's proposal, secondary |V| would denote voicing (and, we add, perhaps also nasality).
33 We refer to Lahiri and Plank (2007) for a review of different views on the relationship between lexical structure and utterance-level prosodic structure. Anderson's view squares with what they refer to as the traditional view, reflected in the work of Abercrombie (1964).
34 Rennison (1987) also uses the aiu-set and a tier-based representational system (without dependency). Goldsmith (1985) adopts some of these elements, while Hyman (2002) uses the unary features low, high, front and round.
35 We note that the notion of gesture was brought to the attention of a general audience in van der Hulst and Smith (1982), in a volume that contained work by many of the later proponents of Feature Geometry.
36 However, early presentations of RcvP unintentionally retain the appearance of constituency and of labels like onset and rhyme as primitives; see the criticism in Anderson (2011c).
37 This extreme position, namely the occurrence of the same element in all groups, was suggested by Petra Kottman, who proposed to use |I| and |U| in all groups, which entailed, of course, a broad set of (phonetically related) interpretations of these two elements.

References

Abercrombie, David. 1964. Syllable quantity and enclitics in English. *In honour of Daniel Jones*, ed. by D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott & J.L.M. Trim, 216–22. London: Longman.

Anderson, John M. 1971. *The grammar of case: Towards a localistic theory*. Cambridge: Cambridge University Press.

Anderson, John M. 1973. *An essay concerning aspect*. The Hague: Mouton.

Anderson, John M. 1986a. Old English ablaut again: The essentially concrete character of dependency phonology. *On language: Rhetorica, phonologica et syntactica: A Festschrift for Robert P. Stockwell from his friends and colleagues*, ed. by C. Duncan-Rose & Th. Vennemann, 156–77. London: Routledge.

Anderson, John M. 1986b. Suprasegmental dependencies. *Dependency and non-linear phonology*, ed. by J. Durand, 55–133. London: Croom Helm.

Anderson, John M. 1987. The tradition of structural analogy. *Language topics: Essays in honour of Michael Halliday*, ed. by R. Steele & T. Threadgold, 33–43. Amsterdam: John Benjamins.

Anderson, John M. 2004. Contrast in phonology, structural analogy, and the interfaces. *Studia Linguistica (Lund)* 58.269–87.

Anderson, John M. 2011a. *The substance of language, volume I: The domain of syntax*. Oxford: Oxford University Press.

Anderson, John M. 2011b. *The substance of language, volume II: Morphology, paradigms, and periphrases*. Oxford: Oxford University Press.

Anderson, John M. 2011c. *The substance of language, volume III: Phonology-syntax analogies*. Oxford: Oxford University Press.

Anderson, John M. 2014. Graphophonology and anachronic phonology: Notes on episodes in the history of pseudo-phonology. *Folia Linguistica Historica* 35.1–53.

Anderson, John M. & Jacques Durand. 1987. Introduction. *Explorations in dependency phonology*, ed. by J.M. Anderson & J. Durand, 1–13. Dordrecht: Foris.

Anderson, John M. & Jacques Durand. 1988. Underspecification in dependency phonology. *Certamen phonologicum*, ed. by P.M. Bertinetto & M. Loporcaro, 1–36. Turin: Rosenberg & Sellier.

Anderson, John M. & Colin J. Ewen. 1980. Introduction: A sketch of dependency phonology. *Studies in dependency phonology*, ed. by J.M. Anderson & C.J. Ewen, 9–40. Ludwigsburg: Strauch.

Anderson, John M. & Colin J. Ewen. 1981. The representation of neutralisation in universal phonology. *Phonologica 1980*, ed. by W.U. Dressler, O.E. Pfeiffer & J.R. Rennison, 15–22. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Anderson, John M. & Colin J. Ewen. 1987. *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Anderson, John M. & Charles Jones. 1972. Three theses concerning phonological representations. *Edinburgh Working Papers in Linguistics* 1.92–115.

Anderson, John M. & Charles Jones. 1974. Three theses concerning phonological representations. *Journal of Linguistics* 10.1–26.

Anderson, Stephen R. 1982. The analysis of French shwa: Or, how to get something for nothing. *Language* 58.534–73.

Archangeli, Diana. 1984. *Underspecification in Yawelmani phonology*. Cambridge, MA: Massachusetts Institute of Technology (Published 1988 by Garland, New York) PhD.

Bao, Zhiming. 1990. *On the nature of tone*. Cambridge, MA: Massachusetts Institute of Technology PhD.

Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language*. Albany, NY: SUNY Press.

Bauer, Laurie. 1994. Structural analogy: An examination of some recent claims. *Studies in Language* 18.1–22.

Blaho, Sylvia. 2008. The syntax of phonology: A radically substance-free approach. Ms, available at http://ling.auf.net/lingBuzz/000672.

Botma, Bert. 2004. *Phonological aspects of nasality: An element-based dependency approach*. Amsterdam: University of Amsterdam PhD.

Botma, Bert. 2009. Transparency in nasal harmony and the limits of reductionism. *Strength relations in phonology*, ed. by K. Nasukawa & P. Backley, 79–112. Berlin and New York: Mouton de Gruyter.

Botma, Bert & Norval Smith. 2006. A dependency account of the fortis-lenis contrast in Cama. *Linguistics in the Netherlands*, ed. by J. van de Weijer & B. Los, 15–27. Amsterdam and Philadelphia: John Benjamins.

Botma, Bert & Norval Smith. 2007. A dependency-based typology of nasalization and voicing phenomena. *Linguistics in the Netherlands*, ed. by B. Los & M. van Koppen, 36–48. Amsterdam and Philadelphia: John Benjamins.

Broadbent, Judith M. 1999. A new approach to the representation of coronal segments. *Issues in phonological structure*, ed. by S.J. Hannahs & M. Davenport, 1–25. Amsterdam and Philadelphia: John Benjamins.

Chomsky, Noam. 1980. On binding. *Linguistic Inquiry* 11.1–46.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clements, George N. 1985. The geometry of phonological features. *Phonology* 2.225–52.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology I*, ed. by J. Kingston & M. Beckman, 283–333. Cambridge: Cambridge University Press.

Clements, George N. & Elisabeth Hume. 1995. The internal organization of speech sounds. *The handbook of phonological theory*, ed. by J. Goldsmith, 245–306. Cambridge, MA: Basil Blackwell.

Colman, Fran. 1987. The phonology and morphology of an Old English digraph: *ie*. *Explorations in dependency phonology*, ed. by J.M. Anderson & J. Durand, 49–77. Dordrecht: Foris.

Colman, Fran. 2005. Old English I-umlaut: A unitary sound change? Dependency, contrast and non-specification. *Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson*, ed. by P. Carr, J. Durand & C.J. Ewen, 31–61. Amsterdam and Philadelphia: John Benjamins.

Davenport, Michael. 1995. The characterization of nasality in dependency phonology. *Leiden in last*, ed. by H. van der Hulst & J. van de Weijer, 89–103. The Hague: Holland Academic Graphics.

Davenport, Michael & Jørgen Staun. 1986. Sequence, segment and configuration: Two problems for dependency phonology. *Dependency and non-linear phonology*, ed. by J. Durand, 135–59. London: Croom Helm.

de Lacy, Paul. 2006. *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press.

den Dikken, Marcel & Harry van der Hulst. 1988. Segmental hierarchitecture. *Features, segmental structure and harmony processes*, ed. by H. van der Hulst & N. Smith, 1–59. Dordrecht: Foris.

Donegan, Patricia J. 1978. *On the natural phonology of vowels*. Columbus, OH: The Ohio State University (Published 1985 by Garland, New York) PhD.

Duanmu, San. 1990. *A formal study of syllable, tone, stress and domain in Chinese languages*. Cambridge, MA: Massachusetts Institute of Technology PhD.

Durand, Jacques. 1988. An exploration of nasality phenomena in Midi French: Dependency phonology and underspecification. *Occasional Papers-University of Essex, Department of Language and Linguistics* 32.30–70.

Ewen, Colin J. 1980. *Aspects of phonological structure, with particular reference to English and Dutch*. Edinburgh: University of Edinburgh PhD.

Ewen, Colin J. 1986. Segmental and suprasegmental structure. *Dependency and non-linear phonology*, ed. by J. Durand, 203–22. London: Croom Helm.

Ewen, Colin J. 1995. Dependency relations in phonology. *The handbook of phonological theory*, ed. by J. Goldsmith, 570–85. Oxford: Blackwell.

Ewen, Colin J. & Harry van der Hulst. 1985. Single-valued features and the nonlinear analysis of vowel harmony. *Linguistics in the Netherlands 1985*, ed. by H. Bennis & F. Beukema, 39–48. Dordrecht: Foris.

Ewen, Colin J. & Harry van der Hulst. 1988. [high], [low] and [back] or [I], [A] and [U]? *Linguistics in the Netherlands 1988*, ed. by P. Coopmans & A. Hulk, 49–58. Dordrecht: Foris.
Firth, J.R. 1948. Sounds and prosodies. *Transactions of the Philological Society* 47.127–52 (Reprinted in F.R. Palmer (ed.), *Prosodic analysis*, 1–26. London: Oxford University Press).
Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8.304–37.
Goldsmith, John. 1976. *Autosegmental phonology*. Cambridge, MA: Massachusetts Institute of Technology PhD.
Goldsmith, John. 1985. Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2.253–75.
Goldsmith, John. 1987. Toward a theory of vowel systems: Parasession on autosegmental and metrical phonology. *Chicago Linguistic Society* 23.116–33.
Gurevich, Naomi. 2011. Lenition. *The Blackwell companion to phonology*, ed. by M. van Oostendorp, C.J. Ewen, E. Hume & K. Rice, 1559–75. London: Wiley-Blackwell.
Hale, Mark & Charles Reiss. 2000. 'Substance abuse' and 'dysfunctionalism': Current trends in phonology. *Linguistic Inquiry* 31.157–69.
Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.
Halle, Morris & Kenneth N. Stevens. 1971. A note on laryngeal features. *Quarterly Progress Report of the MIT Research Laboratory of Electronics* 101.198–213.
Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298.1569–79.
Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20.253–306.
Hays, D.G. 1964. Dependency theory: A formalism and some observations. *Language* 40.511–25.
Heringer, J.T. 1967. Review of Haim Gaifman (1965): Dependency systems and phrase-structure systems. *Information and Control* 8.3.304–37. Ohio State University Research Foundation: Working Papers in Linguistics 1.128–36.
Hjelmslev, Louis. 1953. *Prolegomena to a theory of language*. Baltimore, MD: Waverly Press.
Hume, Elizabeth. 1994. *Front vowels, coronal consonants, and their interaction in nonlinear phonology*. New York: Garland Press.
Itô, Junko & R. Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17.49–73.
Jakobson, Roman. 1968. *Child language, aphasia and phonological universals*. The Hague and Paris: Mouton.
Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.
Kahn, Daniel. 1980. *Syllable-based generalizations in English phonology.* PhD: Massachusetts Institute of Technology (1976). New York: Garland.
Kang, Yongsoon. 1991. Phonology of consonant-vowel interaction with special reference to Korean and Dependency Phonology: University of Illinois at Urbana-Champaign PhD.
Kaye, Jonathan. 1988. The phonologist's dilemma: A game-theoretic approach to phonological debate. *Glow Newsletter* 21.16–19.
Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2.305–28.
Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7.193–231.
Kean, Mary Louise. 1975. *The theory of markedness in generative grammar*. Cambridge, MA: Massachusetts Institute of Technology PhD.
Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. *The Structure of Phonological Representations*, ed. by H. van der Hulst & N. Smith, 131–75. Dordrecht: Foris.
Kula, Nancy C. 2002. *The phonology of verbal derivation in Bemba*. Leiden: University of Leiden PhD.
Lahiri, Aditi & Frans Plank. 2007. On phonological grouping in relation to morphosyntactic grouping. Paper presented to the Annual Meeting of the DGfS, 2007.

Lass, Roger. 1976. *English phonology and phonological theory: Synchronic and diachronic studies*. Cambridge: Cambridge University Press.
Lass, Roger. 1984. *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.
Lass, Roger. 1987. Intradiphthongal dependencies. *Explorations in dependency phonology*, ed. by J.M. Anderson & J. Durand, 109–31. Dordrecht: Foris.
Lass, Roger & John M. Anderson. 1975. *Old English phonology*. Cambridge: Cambridge University Press.
Liljencrants, Johan & Bjorn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48.839–62.
Lombardi, Linda. 1991. *Laryngeal features and laryngeal neutralization*. Amherst, MA: University of Massachusetts Amherst (Published 1994 by Garland, New York) PhD.
Marcus, Solomon. 1967. *Algebraic linguistics: Analytical models*. New York: Academic Press.
McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 43.84–108.
Mester, R. Armin. 1986. *Studies in tier structure*. Amherst, MA: University of Massachusetts at Amherst PhD.
Mester, R. Armin. 1988. Dependent tier ordering and the OCP. *Features, segmental structure and harmony processes*, ed. by H. van der Hulst & N. Smith, 127–44. Dordrecht: Foris.
Ó Dochartaigh, Cathair. 1979. Lenition and dependency phonology. *Éigse* 17.457–94.
Ó Dochartaigh, Cathair. 1980. Aspects of Celtic lenition. *Studies in dependency phonology*, ed. by J.M. Anderson & C.J. Ewen, 103–37. Ludwigsburg: Strauch.
Padgett, Jaye. 2011. Consonant–vowel place feature interactions. *The Blackwell companion to phonology*, ed. by M. van Oostendorp, C.J. Ewen, E. Hume & K. Rice, 1761–86. London: Wiley-Blackwell.
Percival, W. Keith. 1990. Reflections on the history of dependency notions in linguistics. *Historiographia Linguistica* 17.29–47.
Piggott, Glyne L. 1990. *The representation of sonorant features*. Ms., McGill University.
Piggott, Glyne L. 1992. Variability in feature dependency: The case of nasality. *Natural Language & Linguistic Theory* 10.33–77.
Pike, Kenneth L. 1947. *Phonemics*. Ann Arbor: University of Michigan Press.
Prince, Alan & Paul Smolensky. 1993 [2004]. *Optimality theory: Constraint interaction in generative grammar*. London: Blackwell.
Rennison, John R. 1986. On tridirectional feature systems for vowels. *Dependency and non-linear phonology*, ed. by J. Durand, 281–303. London: Croom Helm.
Rennison, John R. 1987. On the vowel harmonies of Koromfe (Burkina Faso, West Africa). *Phonologica 1987*, ed. by W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer & J.R. Rennison, 239–46. Cambridge: Cambridge University Press.
Rennison, John R. 2014. On vowel harmony and vowel reduction: Some observations on canonical shapes of disyllabic nouns in Yukuben, Mòoré and German. *The form of structure, the structure of form: Essays in honor of Jean Lowenstamm*, ed. by S. Bendjaballah, N. Faust, M. Lahrouchi & N. Lampitelli, 37–56. Amsterdam and Philadelphia: John Benjamins.
Riad, Tomas. 2014. *The phonology of Swedish*. Oxford: Oxford University Press.
Ringen, Catherine O. 1978. Another view of the theoretical implications of Hungarian vowel harmony. *Linguistic Inquiry* 9.105–15.
Sagey, Elizabeth C. 1986. *The representation of features and relations in non-linear phonology*. Cambridge, MA: Massachusetts Institute of Technology PhD.
Sagey, Elizabeth C. 1988. Degree of closure in complex segments. *Features, segmental structure and harmony processes*, ed. by H. van der Hulst & N. Smith, 170–208. Dordrecht: Foris.
Sanders, Gerald. 1972. *The simplex feature hypothesis*. Bloomington, IN: Indiana University Linguistics Club.
Schane, Sanford A. 1984. The fundamentals of particle phonology. *Phonology Yearbook* 1.129–55.
Smith, Norval. 1988. Consonant place features. *Features, segmental structure and harmony processes*, ed. by H. van der Hulst & N. Smith, 209–35. Dordrecht: Foris.

Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43.393–436.
Staun, Jørgen. 1985. The representation of irresoluble neutralisations in dependency phonology. *Word* 36.61–76.
Staun, Jørgen. 1996a. On structural analogy. *Word* 47.193–205.
Staun, Jørgen. 1996b. On the location description of consonants. *Travaux du cercle linguistique de Copenhague* 29.1–137.
Steriade, Donca. 1987. Redundant values. *Chicago Linguistic Society* 23.339–62.
Steriade, Donca. 1995. Underspecification and markedness. *The handbook of phonological theory*, ed. by J. Goldsmith, 114–74. London: Blackwell.
Stevens, Kenneth N. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. *Human communication: A unified view*, ed. by E. David & P. Denes, 51–66. New York: McGraw-Hill.
Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
Thráinsson, Hoskuldur. 1978. On the phonology of Icelandic pre-aspiration. *Nordic Journal of Linguistics* 1.3–54.
van der Hulst, Harry. 1988. The dual interpretation of |i|, |u| and |a|. *Proceedings of the North East Linguistic Society (NELS)* 18.208–22.
van der Hulst, Harry. 1989. Atoms of segmental structure: Components, gestures and dependency. *Phonology* 6.253–84.
van der Hulst, Harry. 1994. Radical CV phonology: The locational gesture. *UCL Working Papers in Linguistics* 6.439–77.
van der Hulst, Harry. 1995. Radical CV phonology: The categorial gesture. *Frontiers of phonology: Atoms, structures, derivations*, ed. by J. Durand & F. Katamba, 80–116. London: Longman.
van der Hulst, Harry. 1996. Radical CV phonology: The segment-syllable connection. *Current trends in phonology: Models and methods*, ed. by J. Durand & B. Laks, 333–63. Paris: CNRS/ESRI.
van der Hulst, Harry. 2000. Features, segments and syllables in radical CV phonology. *Phonologica*, ed. by J.R. Rennison & K. Kühnhammer, 89–111. The Hague: Thesus.
van der Hulst, Harry. 2005. The molecular structure of phonological segments. *Headhood, elements, specification and contrastivity*, ed. by P. Carr, J. Durand & C.J. Ewen, 193–234. Amsterdam and Philadelphia: John Benjamins.
van der Hulst, Harry. 2006. Dependency phonology. *The encyclopedia of language and linguistics*, ed. by K. Brown, 220–33. Oxford: Elsevier.
van der Hulst, Harry. 2013. The discoverers of the phoneme. *Oxford handbook of the history of linguistics*, ed. by K. Allan, 167–91. Oxford: Oxford University Press.
van der Hulst, Harry. 2015a. The laryngeal class in RcvP and voice phenomena in Dutch. *Above and beyond the segments*, ed. by J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. Schiller & E. van Zanten, 323–49. Amsterdam and Philadelphia: John Benjamins.
van der Hulst, Harry. 2015b. The opponent principle in RcvP: Binarity in a unary system. *The segment in phonetics and phonology*, ed. by E. Raimy & C. Cairns, 149–79. London: Wiley-Blackwell.
van der Hulst, Harry. 2016a. Monovalent 'features' in phonology. *Language and Linguistics Compass* 10.83–102.
van der Hulst, Harry. 2016b. Phonology. *The Routledge handbook of linguistics*, ed. by K. Allan, 83–103. Abingdon and New York: Routledge.
van der Hulst, Harry. in prep. *Principles of radical cv phonology*. Ms, University of Connecticut.
van der Hulst, Harry. to appear. *Asymmetries in vowel harmony: A representational account*. Oxford: Oxford University Press.
van der Hulst, Harry & Norval Smith. 1982. An overview of autosegmental and metrical phonology. *The structure of phonological representations*, ed. by H. van der Hulst & N. Smith, 1–45. Dordrecht: Foris.
van der Hulst, Harry & Jeroen van de Weijer. 1991. Topics in Turkish phonology. *Turkish linguistics today*, ed. by H.E. Boeschoten & L.T. Verhoeven, 11–59. Leiden: Brill.
van der Hulst, Harry & Jeroen van de Weijer. 1995. Vowel harmony. *The handbook of phonological theory*, ed. by J. Goldsmith, 495–534. Oxford: Basil Blackwell.

van de Weijer, Jeroen. 1992. Basque affricates and the manner-place dependency. *Lingua* 88.129–47.
van de Weijer, Jeroen. 1993. The manner-place dependency in complex segments. *Linguistics* 31.87–110.
van de Weijer, Jeroen. 1996. *Segmental structure and complex segments*. Tübingen: M. Niemeyer.
van Lessen Kloeke, W.U.S. 1982. *Deutsche Phonologie und Morphologie: Merkmale und Markiertheit*. Tübingen: Niemeyer.
van Nice, Kathy. 1991. Hierarchical particle phonology. *Proceedings of LCJL 2*, ed. by J. van Lit, R.H. Mulder & R. Sybesma, 233–44. Leiden: Holland Institute of Generative Linguistics.
Wetzels, W. Leo & Joan Mascaró. 2001. The typology of voicing and devoicing. *Language* 77.207–44.
Wood, Sidney A.J. 1975. Tense and lax vowels: Degree of constriction or pharyngeal volume? *Working Papers of the Phonetics Laboratory, Department of General Linguistics, Lund University* 11.109–33.
Wood, Sidney A.J. 1979. A radiographic analysis of constriction locations for vowels. *Journal of Phonetics* 7.25–43.
Yip, Moira. 1980. *The tonal phonology of Chinese*. Cambridge, MA: Massachusetts Institute of Technology PhD (Published 1990 by Garland, New York).

13

Connectionist approaches to generative phonology*

John Alderete and Paul Tupper

13.1 Introduction

For many, generative phonology is a kind of symbolic logic for sound patterns. Phonological analysis involves the assignment of contrastive sounds to discrete symbols, or defining the phonemic inventory of a language. Further analysis proposes morpho-phonemic and allophonic processes that transform these symbols to new ones in environments that are likewise defined with symbolic sound structure. Phonological structure may also be organized into syllables and other higher level units like prosodic feet that enrich the symbolic phonology with constituency. This article is about looking a layer beneath this symbolic structure and examining how phonological processes can be accounted for at the micro-structure level in connectionist networks. A connectionist network is a web of interconnected micro-processors that is organized and tuned in such a way to implement input–output processes. Individually, the micro-processors do not do anything close to what we think of when we think of phonology. However, when these micro-processors are organized in the appropriate layers, and when the links between them are properly adjusted, the larger network can produce outcomes that have linguistic interpretations and indeed mimic symbol-manipulating phonology. It is possible, therefore, to construct a generative grammar out of a web of interconnected micro-processors, but the resulting model works rather differently than the macro-structure models some have come to take for granted. The goal of this chapter is to illustrate how these connectionist models work precisely and how they have been applied to some of the problems central to phonological theory.

Connectionist phonology did not develop in a vacuum, and so to understand its motivation, it will be useful to understand a bit of its history. Emerging fully in the 1980s, connectionism is an interdisciplinary research program that pursued the idea that cognitive processes arise from brain-like computation and development. Connectionist models are characterized by networks of micro-processors, called units, which are linked together by connections. Information processing in connectionist networks is intended to be a simplified model of how information is processed in human neural networks, where units correspond to individual neurons, and connections correspond to the synapses linking them together. Information in connectionist networks is often distributed across multiple units and passed between layers in parallel, as many researchers believe to be true of brain-like computation. This comparison with the human brain has its limits (see section 13.4.1), but these rather general assumptions have made possible principled explanations of certain key facts about human cognition (see Elman et al. (1996) and McLeod et al. (1998) for useful introductions). The parallel-distributed processing nature of connectionist networks means that they are fault tolerant (resistant to minor damage), content-addressable (memory can be accessed from part of a memory), and well suited for capturing graded categories and similarity structure. The cognitive architecture of connectionist networks supports the formalization of mature cognitive abilities, as well as the development of those abilities in learning through adjustment of the connection weights. This chapter explores how these assumptions apply to phonological systems.

Another important distinction that can be made between connectionist and traditional generative approaches to language is that connectionist approaches are often task oriented (Joanisse 2000). Knowledge of language in connectionist approaches can be built up from the acquisition of several distinct cognitive processes, e.g., the development of sound categories in infants, word learning from age 1 onward, and word production and recognition processes. On this view, knowledge of sound patterns is in a sense emergent from the knowledge acquired in these distinct processes (Plaut & Kello 1999), and so language acquisition in connection science has a focus on the communicative function of language. This approach can be contrasted with the more formal approach to acquisition in generative traditions where language learning is typically modeled as a selection of a particular formal grammar from an inventory of possible grammars. Thus, traditional generative grammar posits a notion of linguistic competence that accounts for all and only the attested linguistic systems, typically vetted by typological study. In learning, children are conjectured to select their grammars from this universal inventory. In contrast, connectionist phonology is often linked to models of linguistic behaviors like speech production (Dell 1986; Dell et al. 1993; Goldrick & Daland 2009), speech perception (Gaskell et al. 1995; McClelland & Elman 1986), and language acquisition (Stemberger 1992). Children do not really select a grammar. Rather, their grammatical knowledge is seen as a by-product of these distinct linguistic behaviors.

While much of connectionist phonology is task oriented and rooted in specific theories of psycholinguistics (see Goldrick (2007) and Stemberger *this volume* for review), there is a significant literature that uses connectionist models to tackle classic problems in generative phonology. Early conceptualization of connectionist phonology in Lakoff (1988, 1993) viewed phonological processes as the work of so-called "constructions", or constraints that hold within or between certain levels of word and sound structure. Lakoff's constructions, which in a sense anticipate later theories of constraint-based phonology, were not explicitly implemented, but his illustrations sketch how the simultaneous application of these constraints could give structure to complex systems of rule interaction and iteration; see Wheeler and Touretzky (1993) for concrete phonological systems using Lakoff's constructions. Another line of research initiated by John Goldsmith (see Goldsmith (1993) for an introduction) grapples with syllable and metrical structure, and illustrates how connectionist networks can address problems of locality and the analysis of graded phonological categories like sonority. We also review connectionist models that were developed to account for other kinds of phonological and morpho-phonological phenomena, including vowel and consonant harmony (Hare 1990; Wayment 2009), disharmony (Alderete et al. 2013), prosodic morphology (Corina 1994), and general models of morpho-phonemic alternations (Gasser & Lee 1990; Hare 1992). Our goal is to explain to the newcomer how these models work and, importantly, how they address problems central to phonology, like the analysis of locality, gradience, opacity, and learnability.

As successful as they are in addressing these problems, it is fair to say that connectionist approaches to generative phonology have been eclipsed by other research trends in phonology, including Optimality Theory (McCarthy & Prince 1995; Prince & Smolensky 1993/2004), Articulatory Phonology (Browman & Goldstein 1989; Browman & Goldstein 1992), and the larger field of laboratory phonology. However, connectionist approaches in fact have considerable common ground with many of these approaches, and other more recent theoretical developments. For example, connectionist phonology shares with Optimality Theory (OT) the idea that phonological systems are constraint based and information is processed in parallel. Indeed, explicit formal parallels are made between the macro-structure of OT grammars and the micro-structure of connectionist networks in Smolensky and Legendre (2006). Connectionism also shares with exemplar phonology (e.g., Pierrehumbert (2003); Wedel (2006)) the assumption that phonological knowledge is built up gradually from experience, and that phonological models must have a robust account of continuous and graded structure (see, e.g., Bybee and McClelland (2005)). Connectionism also provides some of the foundational assumptions of new theories of linguistic knowledge currently in development, including dynamic field theory (Spencer et al. 2009), gradient symbol processing (Smolensky et al. 2014), and information-based theories (Currie Hall 2009). An understanding of connectionist phonology therefore both helps one understand past developments in phonological theory, and it gives a glimpse into the future of phonology.

The rest of this chapter is organized as follows. The next section provides the formal background necessary for understanding how connectionist phonology works. Section 13.3 reviews some core problems in generative phonology and then goes on to survey a range of connectionist models for addressing some of these problems in segmental and prosodic phonology. Section 13.4 goes back over the results of these distinct models and examines the nature of the explanations they offer, as well as pointing out some of the known problems for connectionism. Section 13.5 looks ahead to the future of connectionist phonology.

13.2 Background

Connectionist models differ from most generative models in that they use numerical computation rather than symbol manipulation. Connectionist networks work on vectors and matrices of real numbers and use principles of linear algebra and calculus to produce outcomes of a desired type. But connectionist models of language are still capable of computing the same kinds of functions that generative models do, as shown by many of the examples we survey in section 13.3. At the computational level (Marr 1982), therefore, connectionist models can work on the same kinds of inputs and map them to the same kinds of outputs as symbolic computational models.

Information is processed in connectionist networks by computing the activation states of simple micro-processors called units (see McLeod et al. (1998); McMurray (2000); Smolensky and Legendre (2006); Thomas and McClelland (2008) for a more detailed introduction to connectionism). The activation of a particular unit is a real number that, together with other activation values, can be given a linguistic or psychological interpretation. In other words, the values of a collection of units can represent a linguistic structure, like a syllable or a word. The flow of this information is neurally inspired in the sense that

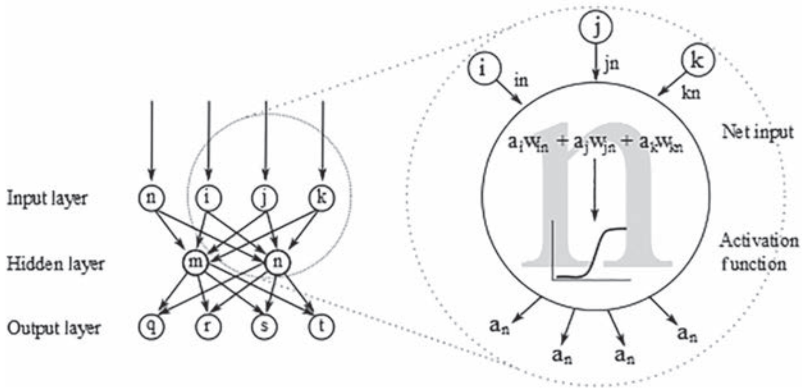


Figure 13.1 Spread of activation in a simple three-layer feed-forward network

the computation of activation states is assumed to be processed in parallel, and the activation state of a particular unit is affected by the activation states of other units connected to it. Figure 13.1 illustrates some of the core ideas of information processing in connectionist networks. In this illustration, units are organized into distinct layers that correspond to different types of representations. For example, the input and output layers encode the structures coming into and out of the network, and these are distinguished from a so-called hidden layer that mediates between the two. Hidden layers are internal representations that can restructure the input in a way that makes possible certain mappings that would not otherwise be possible.

The activation state of any given unit is a function of the weighted sum of the activation states of all units sending information to it. Thus, in the enlarged fragment on the right of Figure 13.1, the activation value of unit *n* is the weighted sum of all activation values feeding into it, i.e., units *i*, *j*, *k*, transformed by a so-called activation function (see the side bar for more definitions and formulas). Concretely, if unit *i* has the activation value of 1.0 and it feeds into *n*, and there is a connection weight of .5 linking the two units, then the *input* from *i* to *n*, that is, the specific contribution from this unit, is .5. The so-called *netinput* into unit *n* is simply the sum of all these inputs. The result of this summation is then transformed by the activation function. The activation function may introduce nonlinearity into the system, increasing the expressive power of the network considerably, and it also forces outputs into a numerical range that supports an interpretation suitable for linguistic analysis. Activation functions are often tailored to the problem at hand, but a common one is a sigmoid logistic function. Summarizing, activation values are the output of the composed function: $f_{\text{activation}}(\text{netinput})$.

To make all of this a bit more concrete, you can compare the flow of information in a connectionist network to the flow of energy on a light board. Imagine, for example, that the units of Figure 13.1 are individual light bulbs that can receive energy to produce light. The greater the energy, the stronger the luminance of the light bulb's output, with potentially infinite shades of luminance. The spread of activation values from the input to hidden, and then hidden to output, can be thought of as the spreading of light patterns generated by the energy surging across the light board, with different light patterns at different states in the spreading action. A particular pattern of light in the input or output can be decoded as a particular representation; it has a meaning in the system.

Panel A. Definitions

Units: micro-processors that receive and pass on information from connected units (artificial neurons)
Weights: real numbers that are co-efficients modulating the spread of activation (artificial synapses)
$input_{ij} = a_i w_{ij}$, or the contribution from a sending unit to a receiving unit
$netinput_i = \sum_j a_j w_{ij}$, or the total contribution to a unit from sending units
Activation function: a function that transforms the output of *netinput* into a bounded range

This metaphor underscores some fundamental differences between connectionist models and symbolic computational models. First, the changing activation patterns involve rather low-level processing of a single currency, activity. Complex patterns and categories are not classified by using different privative elements, but rather with combinations of the activations of simple processing units. Second, since activation values are real numbers, their values can fall anywhere on a continuous dimension, which enables connectionist models to capture graded patterns. This property of connectionist networks makes them well suited for analyzing graded concepts like sonority and similarity in phonology.

As far as giving the input and output layer a coherent interpretation in linguistics, it is typical to distinguish two kinds of representations, local and distributed representations. Local representations are simply vectors of activation values in which a single element in the vector is used to code a specific category. Thus, to encode a particular structure, one unit is active (canonically represented with "1"), and all other units are inactive ("0"). For example, the phoneme [b] can be represented with a vector in which the first element is "1" and all others are "0", so there is a one-to-one relationship between specific elements in the vector and categories in the system. Distributed representations are defined simply as non-local representations: the information coding a particular category is the total combination of activation values in the row vector, and there is no limitation that only one is active and all others are not. Illustrations of these two types of representations are given below.

(1)	Category	Local representation	Distributed representation
	"phoneme [b]"	[1 0 0 0 0 0 0 0]	[1 0 1 1 1 0 1 0]
	"phoneme [d]"	[0 1 0 0 0 0 0 0]	[1 0 0 0 1 0 1 1]
	"phoneme [k]"	[0 0 0 0 0 0 1 0]	[1 1 1 0 0 1 0 1]

While both local and distributed representations are used in language processing, distributed representations are more common in connectionist phonology because they are consistent with feature-based generalization. For example, individual elements in the row vector can correspond to the feature values of distinctive features in phonology, supporting generalizations based on these features. Many networks that try to capture generalizations with phonological features encode feature values in units in some way, e.g., [+voice] might be "1" for a specific unit, and [−voice] "−1" for the same unit.

Another important feature of connectionist networks is how time unfolds. Time is characterized in these models as a series of discrete states or ticks of a connectionist clock that meters the passing of information from unit to unit. Some psycholinguistic models are actually sensitive to time in this way. For example, in Dell's (1986) spreading-interactive model of language production, speech errors occur with a greater frequency in short time intervals (e.g., four ticks in the connectionist clock) as opposed to long time intervals (say, eight time ticks) because, in longer intervals, the network has a chance to settle into the right activation values of intended speech units before they are selected for speech. However, connectionist models of generative grammar in general do not exploit connectionist time in this way. For example, in Goldsmith and Larson's (1990) model of syllabification (illustrated below in section 13.3.2), syllabic role nodes are initialized at a certain level of activation, then activation is sent back and forth from neighboring nodes until the nodes settle into an equilibrium state where they do not change very much. Though the intermediate states of these units do influence the outcome, the model is mainly interested in giving interpretations of the inputs and outputs of the system. In this way, most connectionist generative grammars resemble symbolic generative grammars, as intermediate representations do not have any special interpretation other than they are interim representations that can lead to the correct outcomes.

A different aspect of time, the serial order of elements, has been modeled in connectionism rather differently than generative grammar through the use of a context or state layer. Many of the networks we discuss below are non-sequential in the sense that they do not process sequences as a series of discrete units, but rather an entire structure as a whole. The model of English vowel alternations discussed in section 13.3, for example, processes an entire string of consonants and vowels in tandem. There are important theoretical reasons for this non-sequential property. For example, in psycholinguistic models of speech production (see Stemberger *this volume*), the processing of a particular segment at a particular point in time can look ahead in the speech stream, or even look back, as evidenced by anticipatory speech errors like *[l]eading list* for *reading list*. This fact means that the language processing requires simultaneous access to both future and past speech units, as embodied in a non-sequential network.

However, simultaneous processing of entire strings can be rather limiting for the modeler because it essentially commits the model's representation of processing units to fixed structures. To address this problem, and others, so-called sequential networks were developed. Jordan (1986) developed such a model for sequences of speech segments in order to account for well-known effects of coarticulation (see Lathroum (1989) for a nontechnical explanation of this model), and Elman (1990) uses a similar kind of model for predicting words in full sentences. In these sequential networks, an entire string is not processed as a whole. Rather, the elements of the string are processed one by one, and each element is processed in the context of what has come before. Another aspect of this kind of model that makes it rather different from non-sequential models is that it tends to involve learning an arbitrary association between two structures. Thus, in Jordan's original sequential model for coarticulation, the function computed by the network was to associate a so-called plan representation with a specific sequence of speech sounds. This association is comparable to the Saussurian sign, where the plan can be thought of as a concept (the signified) and the sequence of segments is the phonological word associated with that concept (the signifier). These associations are learned by allowing a feedback loop between the output and input layers. In essence, the hidden layer on each successful pass through the network is allowed to "see" the network's previous output in the form of the state layer. Thus, the first element in a sequence is processed without any prior structure, but the input to subsequent elements is both the plan layer and the context layer. This type of network has been employed in the analysis of harmony phenomena, which is illustrated in section 13.3.5.

Another important component of connectionist networks is how memories are learned. For the most part, when we speak of learning in connectionist models, we mean adjustments to the weight matrix that encodes the associations between two layers. Memories of language are stored in these matrices, and so learning these memories involves changing the values of the specific connection weights in response to experience. In Figure 13.1, for example, the mapping from the hidden to output layers is encoded in an *n* by *m* matrix, where *n* is the number of hidden layer nodes and *m* is the number of output nodes, so two by four, or eight connection weights. Learning the correct mapping from the hidden to output layers involves learning specific values for the eight cells in this matrix. Typically, the values of this matrix are set to 0 or random numbers in the beginning of learning, and then adjusted gradually in response to data. The Delta rule, given below, is a common type of supervised learning rule that makes weight changes in a way that tends to push the network in a direction of making better approximations of the desired mapping. The Delta rule assumes that the learner knows the target activation value of the node *i* and also retains the actual activation value of this node. The difference between the target value and the actual value then is the error. The weight change is the product of the error and the activation value of the input node, scaled by a learning rate ε. Applying this rule iteratively, to all connections in the weight matrix, will gradually reduce the error and therefore improve the network's performance.

(2) Delta rule: $\Delta w_{ij} = [\, a_i \text{ (desired)} - a_i \text{ (observed)}]$ $\quad a_j \quad \varepsilon$
$[\quad$ *error* $\quad]$ $\quad a_{input}$ $\quad$ *learning rate*

Networks with hidden layers require more complex learning rules because we assume that the learner of such a network is not provided with the target activation values of the hidden layer nodes. Backpropagation learning (Rumelhart et al. 1986) involves sending the error signal of an accessible node backward into the network to the inaccessible nodes, in proportion to the weights connecting those nodes, and then updating the deeper connection weights this way (see, e.g., Mitchell (1997) for a good introduction to backpropagation learning). In sum, learning in connectionist networks is error-corrective and modeled as gradual adjustments to weight matrices in response to data.

Finally, connectionist networks can be characterized by their overall organization, and therefore how activity flows through the network. The network illustrated in Figure 13.1 is a feed-forward network in which activity passes successively through a set of layers, similar to the feed-forward nature of phonological derivations in classic generative phonology (though of course the representations are rather different). Connectionist networks sometimes also have feedback loops in which the activation states of some units feed into the units of another layer that has already received input. Such networks are sometimes called recurrent networks. The eight-unit recurrent network developed in McClelland and Rumelhart (1985) to solve certain problems in concept learning is a good example of such a network. Sequential networks, like Jordan's (1986) network developed for coarticulation, are recurrent networks because the output of one pass through the network feeds back into the processing of later units through the context layer. Networks can also be constituted by a single layer of units that interact with each other, as in the single-layer models discussed in sections 13.3.2 and 13.3.3. These networks are designed to model the creation of prosodic structure through competitive inhibition of adjacent units on a single layer of information processing. Thomas and McClelland (2008) is a helpful resource that reviews the historical evolution of these and other neural network architectures.

To summarize the network structures discussed above, the properties below illustrate some of the main ways connectionist networks can be tailored to specific problems, as exemplified in the next section.

(3) Some properties that characterize connectionist networks
  a. Encoding categories in representations: local representations have a one-to-one correspondence between individual categories and units; distributed representations do not
  b. Existence and number of hidden layers: some networks just have an input and an output, while others have one or more hidden layers
  c. Organization of layers: feed-forward networks pass information from one layer to the next, without feedback to prior layers; recurrent networks, like sequential networks, allow feedback
  d. Hidden layer units: the number of units in hidden layers can be important in achieving the correct outcome; generalization of the network to novel stimuli is usually forced by having a fewer number of units than the layer feeding into the hidden layer; memorization typically requires an equal or larger number of hidden layer units
  e. Conception of sequences: non-sequential networks process an entire sequence simultaneously; sequential networks process individual elements of a sequence in their linear order

13.3 Connectionist models of phonology

13.3.1 Classic problems in generative phonology

Let us start with some of the core problems that have been the focus of phonological analysis for some years.

Locality

Phonological processes tend to be local, that is, the target and trigger are generally "close" to each other in some formal sense; even apparent non-local phenomena, like stress and vowel harmony, can be viewed as local with the right representations.

Gradient and scalar phonology

Many phonological phenomena cannot easily be characterized by binary oppositions, and instead generalizations need to be made with respect to scales or continuous dimensions, e.g., sonority, metrical prominence, and similarity.

Opacity

Many phonological effects are not apparent from the surface phonological form. Models of phonology must therefore contend with phonological effects of structure that are hidden from view.

Learning

Phonological inventories and processes are learned effortlessly by small children. Phonological analyses can be evaluated by considering if and how they can be learned.

Few would dispute the importance of having a theory that can give natural solutions to these problems. Below, we flesh out how these problems have been addressed in connectionist models of phonology.

13.3.2 Syllables

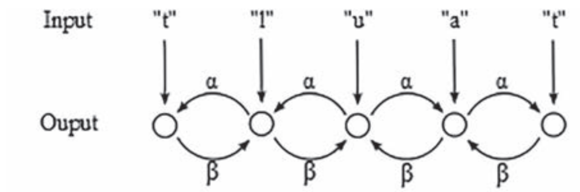
Syllable structure is a good place to start because it is a crucial aspect of any phonological system, and it is a good point of departure for studying connectionist phonology. There are many distinct algorithms for building syllables (Blevins 1995; Itô 1989; Steriade 1982), but, at their heart, syllabification algorithms implement the rather simple idea that syllables are centered over sonority peaks. Thus, syllables are canonically built up in three concrete steps. First, a syllable nucleus is built over high-sonority segments, typically vowels or other sonorants. Syllable onsets are then formed by grabbing a string of rising-sonority consonants and placing them in syllable-initial position. Finally, the residue is dumped into the syllable coda, a process that is subject to certain constraints. Residual material that cannot be put in coda position, for example, may trigger repair rules like epenthesis or deletion.

For many languages, the job of pin-pointing the center of the syllable is a simple matter of finding vowels, which, in turn, triggers the above cascade of operations that fill the Onset and coda positions. There are a number of languages, however, for which the job of assigning syllable roles is rather non-trivial, and following the standard protocol sketched above leads to significant loss of generalization. Dell and Elmedlaoui (1985, 1988, 2002) document such a problem in Tashlhiyt Berber (Afro-Asiatic). Syllables are built in this language by seeking out high-sonority nuclei while, at the same time, requiring all non-initial syllables to have an Onset. This sonority-seeking algorithm, however, is sensitive to eight different levels of sonority, distinguishing two classes of vocoids (low vs. non-low vowels) and six classes of consonants (liquids, nasals, voiced fricatives, voiceless fricatives, voiced stops, voiceless stops). If we follow standard practice of first selecting the correct segment for the nucleus position, this results in eight distinct subroutines, interspersed with Onset formation and other principles dictating how already-syllabified segments must be organized (Goldsmith & Larson 1990; Prince & Smolensky 1993/2004).

An alternative to this approach is to find some natural way of capturing the fact that syllabification is sensitive to the distinct sonority levels. Goldsmith and Larson (1990) provide such an alternative by analyzing sonority as continuous activation values (see also Goldsmith (1992)). In particular, they model the basic competition for the nucleus position in Tashlhiyt Berber as competitive inhibition between adjacent segments. The syllabifications computed by their model are the same as Dell and Elmedlaoui's generative account, and indeed all syllabification systems: it takes a string of segments as input and then returns an assignment of this string to syllabic roles. However, their analysis avoids the need for a myriad of subroutines because their connectionist network captures the effect of sonority with a continuous variable, activity.

Goldsmith and Larson's model is a single-layer network with lateral inhibition of adjacent segments. Figure 13.2 sketches the model and illustrates how it works for the syllabification of the Berber word *tL.wAt* (a Berber place name, capitals are nuclei). Each unit in the model represents a single segment. Different from feed-forward networks, the syllabification algorithm works by changing the states of each unit in the output layer as a function of the influence from its neighbors. As units pass though successive states, they settle into an equilibrium state where their activation values do not change very much (State *n* in Figure 13.2). The resulting output pattern is one with an alternating pattern of "low-high-low-high" activation values that is interpreted linguistically as syllable peaks (high) and margins (low).

Let's flesh out how the specific activation values are calculated in Goldsmith and Larson's model to see how the numerical computation produces this result. Segments are assigned a numerical value for their inherent sonority ranging between 0 and 8 (excluding 6). This value



State 0	0	5	7	8	0
State 1	−3.00	0.80	1.70	7.30	−0.80
State 2	−0.48	4.28	2.54	8.31	−0.73
State 3	−2.57	3.52	1.58	8.18	−0.83
State 4	−2.11	4.30	1.73	8.34	−0.82
⋮					
State n (equilibrium)	−2.59	4.32	1.56	8.34	−0.83

Figure 13.2 Single-layer local network for Berber syllabication, based on Goldsmith and Larson (1990); (α,β) = (.6, .1)

is based on their membership in the eight basic sonority classes and is in principle learnable. Thus, voiced stops like *t* have an inherent sonority of 0, the low vowel *a*, 8, etc. The network is initialized by feeding it these inherent sonority values, in the order they appear in a Berber word, shown in Figure 13.2 at State 0. The activation value of unit *i* at the next state, or the next tick in the connectionist clock, is then calculated as the inherent sonority of *i*, minus the weighted activation value of its neighbor to the left (which it is connected to by weight β) and its right (connected to by weight α) at State *i*. In this illustration (α,β) = (.6, .1), so we can figure out State 1 for *u* by simply subtracting from *u*'s inherent sonority (= 7) the inhibitive value of *l* on its left (= 5 * .1) and *a* on its right (= 8 * .6). Thus, the activation value for the state representing *u* is 7 at State 0, but at State 1 it is: 7 – (.5 + 4.8) = 1.7. The influence from its neighbors has drastically cut the activation value of *u* here, which means its chances of appearing in a syllable peak have just been severely reduced. This same competitive inhibition takes place for all segments simultaneously, with a stronger push downward in activation contributed by the segment on the right, until each unit reaches its equilibrium state.

Goldsmith and Larson's analysis of Berber syllabification gives a very direct analysis of the claim that phonology is local. In their system, which they call "local modeling", segments can only interact with immediately adjacent segments, and yet a global pattern of alternating sonority rises and falls emerges from a series of local struggles for the nucleus position. Also, their analysis captures the graded nature of the sonority scale by assuming sonority is a continuous variable and that all segments compete for a high-sonority position.

Subsequent work has pursued these insights and developed a conceptual framework for relating constraint interaction in connectionist networks to the role of constraints in symbolic theories of grammar like OT. In Legendre et al. (2006), a model similar to Goldsmith and Larson's single-layer model is developed and tested more thoroughly with computational simulations. The Legendre et al. model, dubbed BrbrNet, is similar in having a single layer with competitive inhibition, but with two main differences. First, the inputs to the network are different. Instead of a linear relationship between sonority and activation where both rise by increments of 1, this relationship is exponential in BrbrNet, in particular $2^{son(x)}-1$. Concretely, the initial activation value in Goldsmith and Larson's model grows with increments of 1, as in: 1, 2, 3, . . . 8, but it grows exponentially in BrbrNet, i.e., 1, 3, 7, . . . 255. This exponential growth is argued to be necessary to capture the strict domination structure of OT grammars in the Berber words tested. However, Tupper and Fry (2012) examined the implications of this claim for certain problematic forms and found that in order to give the correct outcomes, this relation needs to be superexponential, a fact that raises problems for the biological plausibility of the model (see below section 13.4.1).

Second, the connection weights between output units are assumed to be symmetric, so the competitive inhibition from the neighbor on the left is the same as the neighbor on the right. In Goldsmith and Larson's model these are two independent parameters, α and β, but Legendre et al. argue that symmetry is necessary in order to implement the principle of harmony maximization. This principle is the analogue in OT grammars to the notion that the winner best satisfies the constraint hierarchy. To understand this point, it is necessary to explain the nature of constraints in connectionism. In connectionist networks, individual connections, or sets of connections, encode constraints (Smolensky 1988). Connectionist constraints may not have as transparent an interpretation as well-known constraints in symbolic phonology, like the ones used in OT. For connectionist constraints, if a connection between two units is positive, the unit sending information tries to put the unit immediately downstream into the same positive state it is in. The constraint is in a sense satisfied if the state of the receiving unit resembles the state of the sending unit. If, on the other hand, the connection is negative, the sending unit tries to put the receiving unit in the opposite state, so negative weights are satisfied by inducing the opposite activation states downstream.

With this definition, we can make parallels between the constraints of connectionist networks and OT constraints. For example, Legendre et al. (2006) show how the negative connections between output nodes in BrbrNet can be compared to Onset in an OT grammar. These negative connections mean that every unit is doing its best to push down the activation value of the segment on its left (and right as well). Interpreted linguistically, this "push down your neighbor on the left" means that segments try to make their neighbors into syllable margins, which is effectively the function of Onset. Getting back to harmony maximization, the gradual changes in the states of the output, as illustrated in Figure 13.2 above, can be seen as a gradual process of working toward a state that better satisfies this "push your neighbor into a margin" goal. Over time, then, the network maximizes harmony (better achievement of constraints), just like OT grammars pick a winner that best satisfies a language-particular constraint hierarchy. While the number of OT analyses for which such parallels have been made is small in number, Legendre et al.'s demonstrations are quite compelling and insightful. A final question raised by the use of symmetric weights is whether onsets should behave just like codas in competitive inhibition. We return to this question in the analysis of French syllabification immediately below.

These analyses of Tashlhiyt Berber illustrate how cumbersome subroutines in a derivational analysis can be avoided, but what does this approach say about other, perhaps more common, syllabification systems? Laks (1995) applied the Goldsmith and Larson model to French syllabification, and showed how subtle facts of French could be modeled with such an account. Additionally, Laks showed that the parameters of the network could be learned through normal processes of error-corrective learning, and the learning results have interesting implications for the nature of the Onset constraints in the model.

Laks constructed a sample of 832 input strings of French and matched the segmental strings with syllabifications based on the intuitions of six native speakers. Laks used a model similar to Larson and Goldsmith's, and modified the parameters of this model in a training phase using an error-corrective technique suggested in Larson (1992). The key differences with the Berber network are that Laks used a different sonority scale tailored to French in the initial parameter settings, and allowed the inherent sonority to be changed in learning. In particular, the initial three-point scale was: consonants = −5, glides = 0, vowels = 5, but after training, a larger set of contrasts was learned that distinguished six important sonority classes necessary for French. Also, Laks distinguished the output links (α,β) for connecting vowels with their neighbors (.5, .5) from those connecting non-vowels (−.5, −.5), and also allowed these to be modified in learning. The use of negative inherent sonority values and connections means that low sonority segments can actually positively contribute to the activation of high-sonority segments like vowels, which is not possible in either BrbrNet or Larson and Goldsmith's original analysis.

Laks presented 20% of the corpus to this algorithm, and then trained the network by allowing the offending segments' inherent sonority to be changed, and the links of this segment and its neighbors to be changed. After training, the mature network was then tested against the rest of the dataset, with near perfect syllabification (99.87% accuracy). While one might object to the "brute force" nature of the learning algorithm, some of the achievements of the learning system are remarkable and worth considering for further research. In particular, the network started with a rather coarse three-way sonority scale, but learned a very natural six-way sonority scale common to many phonological systems. Also, the simulation results show that some segments can have different derived sonority levels based on context. For example, *s* behaves rather differently in /str/ contexts like *apɔstrɔfə* 'apostrophe', from /rst/ contexts, as in *karstə* 'karst', with a much higher derived sonority in coda position. Finally, Laks points out that the network can be naturally extended to account for the gradient intuitions that native speakers seem to have about certain classes, like ambisyllabic consonants, because the derived sonorities in these contexts are less clear-cut than in other contexts.

One important theoretical implication of Laks' learning results is that its accuracy seems to depend on asymmetric parameters for output links. All of the (α,β) values for the six different sonority classes have higher values for the segment on the right than on the left, and these six classes differ in degree of difference between α and β. For example, vowel links are set for (.282, .481) after training, while liquids end up as (−.36, −.3). These results do not show conclusively that the output link parameters must be asymmetric, because it might be possible to just modify inherent sonority in learning. However, they do seem to challenge the claim made in Legendre et al. (2006) that these links are symmetric; see also Touretzky and Wang (1992) on asymmetric connections and directionality in phonology.

13.3.3 Stress

The previous account of syllables showed how connectionist networks can account for global patterns of alternating sonority peaks and falls with simple local interactions. This account made use of a common denominator for assigning syllabic roles, sonority, which is realized as a continuous variable, activity. Goldsmith (1992) extends this approach to stress systems, using essentially the same type of connectionist network, but modeling local competitive inhibition of stress prominence.

Goldsmith's model for stress is again a single-layer network, but instead of representing a sequence of segments, the output layer represents a sequence of syllables. In other words, the network is structured just like the network in Figure 13.2, but with different parameters.

Initial activations are assigned based on the language-particular properties of the stress system (e.g., initial, penultimate, final stress), and then adjacent syllables compete with each other for prominent syllable status. The table in (4) illustrates how this competition accounts for a stress system with initial stress and alternating secondary stresses. At State 1, the first unit, representing the first syllable, is assigned an initial jolt of activation, but all other units (= syllables) have no activation. At the next state, the activation of the second unit is $0.0 + (1 * -.7) = -0.7$. This in turn leads to a positive contribution of .14 to the first unit at State 3 because the weight connecting a unit to its neighbor on its left is −.02, so $a_1 = 1 + (-0.2 * -.70) = 1.14$. The local influences on neighboring syllables spread through the word until again the layer settles into an equilibrium state where the unit activation values do not change very much. The resulting pattern, shown at the bottom of (4), is then interpreted as the surface stress pattern, perhaps after transforming the numbers to more discrete values.

(4) State changes for stress system with initial main stress, alternating secondary parameters: $K(1) = 1.0$, $K(i) = 0.0$, $(\alpha,\beta) = (-0.2, -0.7)$

	σ_1	σ_2	σ_3	σ_4	σ_5
State 1	1				
State 2	1	−.70			
State 3	1.14	−.70	.49		
State 4	1.14	−.90	.49	−.34	
State 5	1.18	−.98	.70	−.34	.24
State 6	1.18	−.98	.70	−.54	.24
State 7	1.19	−.98	.78	−.54	.37
. . .					
State *n*	1.20	−1.01	.84	−.69	.48

The single-layer model illustrated here has the appearance of a single linguistic representation, either a single string of segments or syllables. As a result, it may seem to lack the derivational steps that are necessary for some kinds of phonological effects. In stress systems, for example, stress assigned on a prior cycle can account for a stress lapse that is retained in a later stage of the word, even though the stress that led to the lapse has been lost. This kind of opaque interaction is argued to require derivational stages in Cohn's (1989) analysis of Indonesian stress (cf. Cohn and McCarthy (1998)), as illustrated below for six-syllable words. The lack of a secondary stress on the third syllable in (5b) is due to the presence of a penultimate stress at the stage when stress is assigned to the five-syllable stem, which is later de-stressed because of word-level penultimate stress in the suffixed form.

(5) Six-syllable words in Indonesian: monomorphemic vs. polymorphemic
   a. ò o ò o ó o
   b. ò o o̲ o ó + o

Goldsmith (1992) shows how this kind of "hidden structure" can be accounted for more directly with the subtle dynamics of his model by simply assigning initial activation values of 1.0 to the penultimate syllable of both the stem and the word. In particular, he shows how the initial high activation of the fourth syllable in (5b) pushes down the activation of the third syllable, accounting for its lack of stress. But, at the same time, the fifth syllable pushes down the activation of the fourth syllable, resulting in it ultimately having activation value consistent with the lack of stress. In a sense, the initial states of the fourth syllable act like an intermediate representation, triggering an effect but ultimately fading away as the layer reaches equilibrium. However, the analysis does not require distinct levels of representation, as the cyclic account does, and the opacity effect is produced with surprisingly little theoretical machinery.

Prince (1993) gives a formal analysis of Goldsmith's connectionist model for stress, and probes its typological consequences. In general, it is treated as a predictive model, like other generative models, and it is evaluated based on how well it accounts for all and only known stress systems. On the positive side, Prince points out that the model accounts for all the stress patterns in Prince (1983), an authoritative reference on certain kinds of stress patterns. It is also successful in accounting for stress window effects. Because of built-in rhythmic alternation, stress is forced to fall within three units of the end of a string, which accounts for well-known cases like Spanish that limit stress to three syllables from the end of a word. However, it appears that the model also over-generates, as it predicts many patterns that do not appear to exist in the world's languages. For example, if an input has two prominent syllables, the model can output an alternating string of stress that begins at one prominent input and ends at the other end, with the rest of the word unstressed (cf. Indonesian polymorphemic words above). Another example is a system where medial syllables are stressed as a rule, rather than the universal pattern of main stress aligning with an edge. It is clear from Prince's investigation that some "pathological systems" are predicted, but one might also reply that Goldsmith's core model is largely unconstrained and predicts a number of well-attested systems with just a handful of free parameters. Perhaps limitations on the range of initial activations (as argued in Idsardi (1992) for lexical stress systems) would produce a better goodness of fit between predicted and attested cases.

13.3.4 Segmental mappings in morpho-phonemics

Moving down to the segmental level, any model of phonology will have to contend with morpho-phonemics. Morpho-phonemic processes can instantiate automatic phonological changes, e.g., devoicing in English plurals, or non-automatic changes in the sense that they are associated with particular constructions or lexical strata, like the ablaut alternations in English strong verbs. We illustrate a model of morpho-phonemics using a non-automatic process in English, because, as a result of the legacy of the English past tense debate (see McClelland and Rumelhart (1986); Pinker and Prince (1988) *et seq.*), fully implemented models of non-automatic processes are far more prevalent. The underlying mechanisms of spreading activation are the same as those used for automatic morpho-phonemic processes, so the analyses of these non-automatic processes extend to simpler automatic processes. The model developed in Plunkett and Marchman (1991, 1993) for the vowel changes in English past tense provides a representative example of this kind of system. Like Plunkett and Marchman's model, a number of connectionist models have been developed to largely address problems in morphology, but in the process, account for non-trivial phonological alternations (Gasser & Lee 1990; Hahn & Nakisa 2000; Hare et al. 1995; Plunkett & Nakisa 1997); see Stemberger *this volume* for a review of these models and other models, and Anttila (2002) and Inkelas et al. (1997) for discussion of the nature of these problems in symbolic phonology.

Plunkett and Marchman's model is a feed-forward, multi-layered, non-sequential network that uses distributed representations to encode the phonological structure of present and past tense stems. Figure 13.3 illustrates these basic properties. The input to the system is a distributed representation of a three-segment present tense form. Each segment is encoded

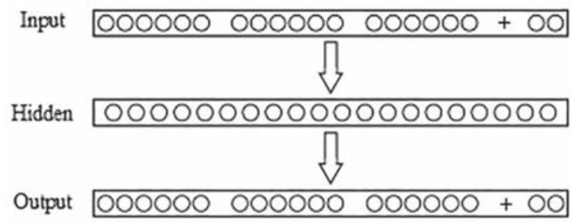


Figure 13.3 Three-layer feed-forward network for English past tense, based on Plunkett and Marchman (1991, 1993)

as a sequence of 0s and 1s in a six-node sequence, and the values for these nodes correspond to values of phonological features necessary to uniquely distinguish segments (i.e., features coding major class, voicing, place, and manner). Thus, the sound *p* is represented as [0 1 1 1 1 1], which contrasts with *b* in the second slot reserved for voicing information: [0 0 1 1 1 1]. Three segment inputs and outputs therefore have 18 nodes for the stem (3 * 6) and they also have a final two-node sequence for encoding the allomorphs of the past tense suffix, i.e., *-t, -d, -əd* and -∅ (i.e., the absence of a suffix, as in strong verbs). The function computed by the network is therefore one that maps present tense verbs to their correct past tense forms, including modifying the vowels in irregulars that exhibit ablaut alternations. This mapping is achieved by spreading activation from the input to hidden layer consisting of 20 nodes, and then from the hidden layer to the output layer. It is thus feed-forward because activation values spread from one layer to the next in a uniform direction. It is also non-sequential because the network has no conception of how the pronunciation of segments unfold in time. The first segment is simply the one represented with the first six nodes, the second the next six nodes, etc., and all segments are presented to the network simultaneously as the input. Finally, the network has three layers, including a hidden layer that can restructure the input in a way that makes possible certain associations and generalizations. Plunkett and Marchman (1991) compared this three-layer network to a simpler one with just two layers (after McClelland and Rumelhart's (1986) original model for English) and found that this hidden layer was indeed necessary to capture the facts of English.

Examining the activation dynamics in an actual word is useful to explain how the network works. To do English morpho-phonemics, the network must learn both the vowel changes in irregular verbs and the lack of vowel change in regulars. Concretely, it must learn that the past of *meet* [mit] is *met* [mɛt], but also that a suffix is required and the vowels do not change in *cheat* → *cheated*. The input–output pairing in (6) illustrates what this means concretely for *meet–met*. Thus, in the environment *m__t*, the 9th, 10th, and 12th node must change from a "1" to a "0", and everything else must stay the same, but this vowel change must not happen in the *ch__t* environment. The specific model parameters are not given in Plunkett and Marchman (1991), but we know that the input will be restructured in the hidden layer in such a way that it can be classified as an irregular verb and that the combined input-to-hidden and hidden-to-output mappings will change the "1"s to "0"s in the right slots.

(6)	Input:	[m]	[i]	[t]	∅
		010011	111111	001110	+ 00
	Output:	[m]	[ɛ]	[t]	∅
		010011	110010	001110	+ 00

An aside about this network is that its task orientation makes it a little different than typical generative models that map abstract underlying representations onto surface forms. The network simply learns to associate actual present tense forms to actual past tense forms. Though the network does use a hidden layer, which might be compared to something like an intermediate representation (with important differences), the main point here is that the model does not assume a native speaker has abstract information about the input of the present and past tense forms. Learning English morphology is about learning the association between two existing words (see recent discussion in the generative literature, e.g., Albright (2002), also casting doubt on the role of underlying representations).

Another aspect of the hidden layer worth commenting on is the number of hidden layer nodes. Plunkett and Marchman (1991) varied the number of hidden layer units from ten to 120 units and found that 20 units was a good compromise between an attempt to optimize performance and maximize the generalization properties of the network. This is likely due to the fact that the English past tense exhibits both sound based generalizations, i.e., the family resemblances within strong verbs, and many exceptions. The network therefore needs a sufficient number of units for coding the exceptional sound patterns. Simpler, more systematic phonology, like final devoicing, however, can be coded with far fewer units because the associations between the natural classes of the input and output are cleaner. Hidden layers with far fewer nodes than the nodes of the inputs and outputs are often used as bottlenecks that force generalizations, whereas a large number of nodes permits item-level associations akin to rote memorization. Connectionist modelers therefore sometimes have to experiment with the number of hidden layer nodes to find the right range suitable for their data. While it is sometimes argued that language-particular and phenomenon-specific hidden layers are descriptive in nature and challenge universal conceptions of the cognitive architecture, the specific number of hidden layer nodes is in principle learnable through mechanisms of sprouting and pruning nodes (Fahlman & Lebiere 1990; LeCun et al. 1990), so this argument is more complex and requires further investigation.

13.3.5 Assimilation and dissimilation

In sections 13.3.2 and 13.3.3 we have examined how continua like sonority and stress prominence are captured in connectionist models. Another kind of continuous structure that is an important factor in phonological processes is phonological similarity. The similarity of two segments is rather important in the analysis of segmental phonological processes. For example, many studies of dissimilation have shown how graded categories of similarity are necessary for capturing place co-occurrence restrictions (Frisch 1996; Pierrehumbert 1993). Similarity is also crucial to the analysis of harmony rules, as the activity of a harmony rule is often predicated on some kind of shared feature structure. Connectionist networks are good at capturing graded categories of similarity because distributed representations of activation patterns are sensitive to similarity structure that is not easily captured in symbolic models. We review below some connectionist analyses of harmony and disharmony phenomena that capitalize on these features of connectionist networks.

Many vowel harmony rules only apply when the target and trigger are sufficiently similar in the phonological feature space. Building on the insights of Jordan (1986) for coarticulation (see section 2), Hare (1992) builds a connectionist model of Hungarian vowel harmony

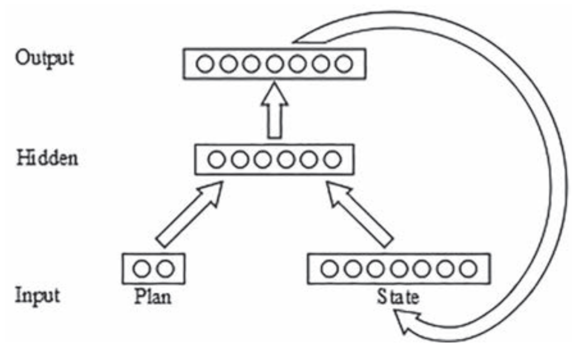


Figure 13.4 Sequential network of Hare (1992) for vowel harmony

specifically designed to address this problem. The analysis requires two key assumptions: (i) that certain nodes can be unspecified for an activation value and thus acquire its value from the nodes representing neighboring segments, and (ii) activation of a current layer is influenced by the output on a prior cycle. Hare's model accounts for the second assumption with a sequential model in which the output of the model, a distributed representation of a vowel feature matrix, cycles back to a state layer, which is then fed as input for the processing of the next vowel (Figure 13.4).

Let's first flesh out how exactly the model works as a model of vowel harmony, and then return to the issue of capturing the similarity effect. The larger function computed by the network is a mapping of a plan input to a sequence of vowel outputs that together constitute a string of vowels. The plan is an arbitrary vector of activation values that functions something like the linguistic concept that the vowel sequence corresponds to. In other words, if we are interested in generating the surface word *CaCi-Ce*, the units associated with the plan layer are just an arbitrary set of activation values that triggers that sequence of vowels. This model is really different, therefore, from feed-forward models like Plunkett and Marchman's model of English because it is a mapping from a linguistic concept to a phonological form, not one phonological form to another form. The output at each step is a seven-unit distributed representation of a vowel where the activation values of each node correspond to traditional phonological features for vowels (i.e., backness, height, roundness, sonority). As a sequential model, the complete phonological form is generated by successive cycles through the network, where each distributed representation of a vowel is both the output representation of a vowel and the context for the next vowel. Thus, the activation vector of each vowel output is fed back into the network as the state layer. For example, the input for the second cycle, which generates the second vowel, is both the plan input and the state input, which is a kind of memory buffer of the vowel pattern just generated, i.e., the first vowel. The associations between the plan and the sequence of vowels are learned through error-corrective backpropagation learning (see section 13.2).

The simulation results sketched below in (7) show how Hare's model captures the similarity effect for some key examples. The seven element vectors on the right show how the network encodes ideal vowel features. These are target activation values that the network is trained on – the actual values produced by the network are close approximations of these. The network is designed to model [back] harmony, so the first element in the vector under the [back] column below is unspecified in the last vowel of the sequence. In this case, the final output determines the a/e value of the suffix *-nAk*, which in Hungarian marks the dative.

There is no target value for [back] in the last cycle through the network, so it gets its value (underlined below) from the state layer. Which prior vowel colors this final vowel is a matter of phonological similarity, computed as a function of the shared non-back features (the shared features are boxed below). When the closest vowel, V2, is more similar than other vowels, its values are carried over in the final run through the network, as shown in (7a). Here, the actual value for the unit associated with [back] in V3 is .86, but a rounding-up procedure enables this to be interpreted as a "1", which is the value for [+back]. If, on the other hand, a vowel in a non-adjacent syllable is more similar than the local vowel in V2 position, the suffix vowel harmonizes with the more similar but distant vowel. In (7b), *a* shares all its features with the suffix vowel, while *i* shares virtually no features, so *a* is the trigger. Hare's analysis of the similarity effect thus accounts for a basic fact of the system, which is that *i* is transparent, i.e., not a trigger in vowel harmony. A curious fact of Hungarian, however, is that if the suffix vowel is preceded by two transparent *i* vowels, a front vowel does trigger harmony. This fact is also accounted for in Hare's analysis, because the network can look back two syllables for a similar non-adjacent trigger, but no further than this, as demonstrated in (7c).

(7) Similarity effects in Hungarian vowel harmony

			Target vowel vectors						
	Position	V	back	height			rd	son	
a. Local trigger V2 more similar than V1	V1	ü	0	1	1	1	1	0	0
	V2	o	1	0	0	1	1	0	1
e.g., *pügo-nak*	V3	a/e →a	.86	0	0	0	0	1	1
b. Nonlocal trigger V1 more similar than V2	V1	a	1	0	0	0	1	1	1
	V2	i	0	1	1	0	0	0	0
e.g., *taxi-nak*	V3	a/e →a	.89	0	0	0	1	1	1
c. Two syllable threshold	V1	a	1	0	0	0	0	1	1
e.g., *anali:zis-nek*	V2	a	1	0	0	0	0	1	1
	V3	i	0	1	1	1	0	0	0
	V4	i	0	1	1	1	0	0	0
	V5	a/e →e	.08	0	0	0	0	1	1

In sum, the sequential network captures the similarity effect, which is both the result of the activation patterns of prior vowel outputs and the formal limits on retaining the memory of these prior activation patterns.

A somewhat different approach is taken in Wayment (2009) to the similarity effect on phonological processes, which illustrates some additional theoretical assumptions. Consonant harmony, while relatively rare in the world's languages, exhibits the same kind of phonological similarity requirement on the target and trigger as vowel harmony. Thus, in Ineseño Chumash, consonants that share the features [+continuant, coronal] agree in the feature [anterior], e.g., /k-su-ʃojin/ → *k-ʃu-ʃojin* 'I darken it' (Hansson 2001). Like Hare's approach, Wayment (2009) captures the similarity structure of target and trigger in a connectionist network, but the specific mechanism is rather different. Instead of implementing time as a sequence of cycles through a recurrent network, time is captured in a time vector, a specific string of units dedicated to representing the position of a segment in a string that is distinct from the units used to code features. The vector for feature values and the time vector are then combined through a process of filler-role binding that makes use of tensor product representations (Smolensky 2006). The feature vector encodes the filler of a segment, i.e., what segment it is, and the time vector encodes its role, or where it appears in the string. The entire segment, in a particular position, is encoded with the tensor product of the two vectors, whose dimension is the product of the two vectors. Wayment convincingly shows how phonological similarity can be captured in a single-layer network with these filler/role representations (in particular, a Hopfield network), and harmony can be predicated on this similarity together with locality. Wayment further shows how the constraints of his network resemble the properties of a set of attraction constraints in OT (Burzio 2002a; Burzio 2002b), illustrating another parallel between the micro-structure of connectionist networks and the macro-structure of OT.

Similarity and gradience have also been the focus of many investigations of dissimilatory phenomena, i.e., dissimilation processes where two similar sounds become less alike, or static root co-occurrence restrictions that achieve the same effect. While some patterns of dissimilation are nearly categorical, dissimilation tends to be a statistical fact of the lexicon and its strength scales with phonological similarity of segments. For example, in Arabic, restrictions against homorganic consonants are stronger for consonants that share more features. As shown in Frisch et al. (2004), phonological similarity between two consonants, established through a metric of shared phonological features, negatively correlates with the representation of the pair in the lexicon. This statistical effect is clearly psychologically real, because native speakers are sensitive to the similarity avoidance effect when they rate nonsense words (Frisch & Zawaydeh 2001). Capitalizing on the ability of connectionist networks to capture gradient effects such as this, Alderete et al. (2013) constructed a connectionist grammar for assessing Arabic roots and analyzed its properties.

Alderete et al.'s model is a non-sequential multi-layer network that takes distributed representations of Arabic triliterals as inputs and outputs a value that assesses the triliteral on a continuous scale from −1 to 1. This network functions differently than other networks, as it does not transform one string into another (Plunkett and Marchman's model of English), and it does not associate a plan with a phonological form (Hare's model of Hungarian). Rather, it functions something like a grammar that takes inputs and assesses them for their overall grammaticality (see Ramsey et al. (1990) for a similar use of output nodes in syntax). In particular, the input has 51 units, or three sets of 17 units, where units encode the feature specifications of the three consonants. The network uses the feature assumptions of Frisch et al. (2004), which is essentially the feature system of Clements and Hume (1995), adapted for Arabic. The activation values of the input spread to a hidden layer of five nodes and then onto the output node responsible for assessing the input. The connection weights between the output node and the hidden layer, and the hidden layer and the input, were trained on a comprehensive sample of Arabic roots, where the connections were gradually adjusted so that attested roots caused the network to produce a value close to "1", and unattested roots a "−1". The trained network was shown to capture the effects of similarity, both in a comprehensive test of nonsense words and judgement data from native speakers. In particular, the values for the output node were compared with the human judgement data classifying Arabic roots in Frisch and Zawaydeh (2001), and the network accounted for the same effect of phonological similarity on grammaticality found in this study.

Alderete et al. also scrutinized the internal workings of the network to see how it relates to macro-structure analyses of dissimilation. In particular, using certain statistical techniques, they examined the behavior of each hidden layer node in the trained network to see how it classified the data. In each of three trials, they found that the functions computed by the hidden layer nodes corresponded to a known set of constraints in constraint-based phonology, Obligatory Contour Constraints for place features (Myers 1997; Suzuki 1998). For example, one hidden node functions like the OCP for [pharyngeal] specifications, another hidden layer node for OCP[dorsal], etc. Moreover, the network was shown to capture the differing magnitudes of these distinct constraints and their idiosyncratic exceptions. This example therefore shows, like BrbrNet, that connectionist networks can closely parallel the effects of known phonological constraints.

13.3.6 Other phenomena

The above survey is by no means exhaustive of the types of phonological processes that connectionist networks have been designed to capture, but it is a good overview of the types of models employed. We summarize briefly some additional phenomena that connectionist models have been built for, and also sketch a few of the problems that have not been solved yet.

From the discussion of connectionist approaches to syllabification and stress, one might form the impression that prosodic constituents themselves are not necessary. Segments and syllables are organized into larger groups centered over peaks of different kinds, but there is no need to invoke the category of a syllable or metrical foot. Many phonological and morphological phenomena do seem to require reference to prosodic structure, and one well-known case is prosodic morphology (McCarthy & Prince 1986; McCarthy & Prince 1993). In prosodic morphology, morphemes are shaped by language-particular prosodic units, and part of the analysis has to determine just how the shape facts are predicted. Corina (1994) investigated this problem in Ilokano reduplication, testing to see if a particular type of connectionist network could induce the CVC shape of Ilokano reduplicative prefixes. Corina built a sequential network (as in Hare's model above) that produced a segment-by-segment output of the desired phonological form. The input to the model was a local representation that combined semantic information and a distributed representation of either a plain form or reduplicated form. After training, the network was found to make many errors, and so it cannot be said to fully account for the facts. However, the network did learn the gross CVC pattern, which the author attributes to the network's sensitivity to the sonority of the input segments (a structure that was implicitly encoded in the segments) to infer larger prosodic structure. One limitation of the model, shared with Hare's model of vowel harmony, is that the network only has memory of the two previous segments it has generated. This is a general problem of the sequential networks based on Jordan's (1986) original design, so perhaps the deeper memory into prior structure allowed in newer models like Elman's (1990) simple recurrent networks would help improve performance.

Connectionist models have also been developed to account for other phonological processes like epenthesis, deletion, and devoicing (Gasser & Lee 1990; Hare 1992; Hare et al. 1989). It is fair to say, however, that the difficulties of implementing connectionist networks have hampered progress in covering more phonological ground. Well-known segmental processes like palatalization and laryngeal alternations have not really been studied, and tone has also been largely ignored. While initial conceptions of connectionist phonology had a broad vision of grappling with complex rule systems and interaction among various linguistic levels (Lakoff 1993; Wheeler & Touretzky 1993), and while some progress has been made on focused problems (Touretzky & Wheeler 1990a; Touretzky & Wheeler 1990b; Touretzky & Wheeler 1990c; Touretzky & Wheeler 1991), we do not know of any implemented connectionist analyses that approach anything like the rich rule complexity found in cases like Mohawk phonology (Halle & Clements 1983). Another lacuna seems to be the Elsewhere Principle, the idea that specific processes take precedence over more general ones (Kiparsky 1973), though see Tabor et al. (2013) for an analysis of the emergence of the Elsewhere Principle in a classification task. Perhaps one avenue of future exploration is to model the gradual phonological processing of harmonic serialism (McCarthy 2000) at the micro-structure level. Indeed, the basic conception of harmonic serialism, that phonological processes proceed step-by-step and incrementally maximize harmony with respect to a constraint ranking, is rather parallel to the workings of recurrent networks (see discussion in 13.3.2). In sum, connectionist approaches have not fully addressed some of the problems that phonologists are interested in, but there are some tractable ideas that may help progress towards meeting this goal.

13.4 Explanations, and challenges to them

To put the models reviewed above in a broader perspective, we reexamine some of the explanations they give to problems in phonology, and also flesh out some of the challenges still faced by this approach.

13.4.1 Biological plausibility

Some of the initial impetus for connectionist research is the idea that it implements cognitive processes with brain-like computation. Surely, the principles of parallel processing and distributed representation have brought the program a big leap forward in this regard, but many issues remain with the biological plausibility of connectionist networks. The first is based on the analogy between connectionist nodes (or units) and human neurons. Connectionist units have continuous activations, but actual neurons are different from these units in that they are either firing or not. The firing of a neuron occurs at effectively a single instant in time, and then the neuron goes into a refractory period before it can fire again. Some psycholinguistic models actually include a refractory period, like the resetting of activation values to zero in spreading-interactive models of speech production (Dell 1986; Stemberger 2009). But even in these models, firing-refractory states are not broadly invoked across the board, and most linguistic models do not employ such a mechanism. Another problem is that neurons are very sparsely connected, but connectionist models tend to have a rich set of interconnections.

In order to interpret connectionist simulations as neurological in nature, it is necessary to interpret activation as firing rate, i.e., the number of fires per second, and each unit as representing the aggregation of many neurons. Thus, the activation of a single unit can be thought of as corresponding to the average firing rates of many neurons (Dayan & Abbott 2001). This interpretation puts strong constraints on dynamic connectionist networks if we want them to be biologically plausible. For example, the shortest time interval between two firings of the same neuron is on the order of one millisecond. It is therefore unreasonable to expect that significant changes in the firing rate of a neuron (i.e., an activation of a unit) can significantly change over a shorter time interval than that. For many linguistic tasks that take place on the time scale of seconds, the ratio of time scales between the fastest process in the network to the slowest process can be at most 10^5. Any network that utilizes a greater range of time scales is not biologically plausible. As an example, Tupper and Fry (2012) show that connectionist networks implementing OT-style constraint systems, such as BrbrNet (see section 13.3.2), require a greater range of time scales than this to function properly. Furthermore, this syllabification system was rather simple, involving only a handful of well-formedness constraints. The time scale problem becomes even more difficult as the number of constraints increases.

Another difficulty for connectionist modeling of language has to do with training connectionist networks. Hebbian learning as described above and used explicitly by some linguistic models (e.g., Wayment's (2009) model for consonant harmony) has broad empirical support as a neural-level model of learning. However, Hebbian learning cannot effectively train connectionist networks with hidden layers, and hidden layers have been shown to be crucial to the success of many connectionist models, like Plunkett and Marchman's model of English morpho-phonology (see section 13.3.4). As explained in section 13.2, models with hidden layers require backpropagation of the error signal. However, backpropagation as it is usually implemented is unlikely to occur in human neural networks (but see O'Reilly (1996); Hinton (2016) for some possibilities). In sum, before the connectionist analyses fleshed out here can cash out on the biological plausibility argument, a serious reexamination of the relation between units and neurons, as well as learning, must take place. It should be emphasized, however, that the time scale problem and learning issues are not unique to connectionist models. Connectionism simply makes specific assumptions, some of which directly address the mind–brain problem, and these assumptions lead to difficult questions about how to interpret signal processing in explicit biological models of human neural networks.

13.4.2 Learning

Another attractive aspect of connectionist modeling is its basic approach to learning. There are well-studied algorithms for training connectionist networks that, once set to initial random weights, do surprisingly well at both modeling known patterns and generalizing to new ones, as illustrated with many of the above examples in section 13.3. Furthermore, the gradual adjustment of connection weights achieved by these algorithms can be linked in natural ways to the processes that underlie language development. For example, the gradual accumulation of phonological patterns can be seen as a natural by-product of word learning. Some of the models reviewed in section 13.3 seem successful in this kind of task-oriented approach to learning phonology. For example, Hare's model of learning vowel harmony learns associations between plans and pronunciations, i.e., the relation between concepts and phonological structure, which is essentially word learning. Wayment's model of learning consonant harmony is likewise consistent with word learning, and has the added bonus that it relies only on Hebbian learning. Finally, recent work has also shown how the identity of phonological well-formedness constraints can be learned in connectionist networks (Alderete et al. 2013; cf. Hayes & Wilson 2008).

However, some problems addressed by connectionist models have not really been completely solved. Laks' model of learning French syllables has incredible accuracy, but one might object that the learning algorithm it uses is too brute a force and allows adjustment of too many free parameters (i.e., the initial activation, and connection weights both to and from all neighbors). This in turn means that it can learn unattested linguistic patterns. Alderete et al.'s approach to learning the OCP also performs well, but the mappings achieved by the model cannot as yet be thought of as a model of production or perception, so the network behavior does not yet have a natural psycholinguistic interpretation. Examination of the errors produced by connectionist networks can also weigh in on how well the network parallels language development (see, e.g., Plunkett (1995)), and further study of phonological development in connectionist networks must attend to this.

13.4.3 Gradience and scales

One of the clear advantages of connectionist approaches to phonology is that they are naturally gradient, so they give direct analyses of graded phonology and scales. The units that make up layers and the connection weights themselves have continuous values, which permit infinite shades of grey in terms of capturing points on a linguistic dimension. The examples in section 13.3 illustrated the importance of graded categories in many domains, from suprasegmentals (sonority-based syllabification and stress) to segmental phonology (assimilation and dissimilation). These analyses lead to two new questions. The suprasegmental analyses are of interest because they seem to obviate the need for phonological constituency, at least for these select phenomena. We ascribe "peak" and "margin" categories to the higher and lower sonority elements in Berber syllabification models, but this does not mean the segments should be interpreted as forming syllables. These analyses can thus account for the variant realizations of segments, like the difference between a glide and vowel in Tashlhiyt Berber, but the analyses themselves do not involve constituents. One might reasonably ask, then, if phonology needs these constituents at all? It seems unlikely that all of the phenomena traditionally ascribed to prosodic units can be modeled with strictly local interaction. There are just too many phonological processes that depend on the foot and the syllable, and they do not seem easily accounted for with a kind of alignment of high or low activation values. How would laryngeal neutralization in codas be approached or spreading rules that make reference to foot structure? It seems therefore that some mechanism for positing prosodic constituency, and even feature geometry, seems necessary, and the tensor product representations developed in Smolensky (2006) (see section 13.3.5) are suitable to this task.

Another issue is how other known scalar phenomena might be treated in connectionist networks. While modern phonology tends to break sound structure into a set of binary oppositions, a number of phonological processes seem to be sensitive to intrinsic scales that are not easily captured by this simple system of contrast (Foley 1970; Gnanadesikan 1997), like Gnanadesikan's inherent voicing scale: voiceless obstruent < voiced obstruent < sonorant. Perhaps these scales, which are often ternary in nature, could be captured in connectionist grammars. A fundamental distinction is made between "adjacent" and "non-adjacent" elements on these scales, and if a natural linguistic dimension could be established that ties all elements on the scale together, then continuous activation values would be suitable to this kind of problem. In other words, the approach to scales is not limited to phonological similarity and sonority.

13.4.4 Algebraic phonology

One problem that plagues many connectionist networks is that they cannot easily instantiate variables. To make this problem clear, it is necessary to distinguish a certain type of connectionism, namely associationism, from other types, like the models found in Smolensky and Legendre (2006) and Eliasmith (2013), which are in general quite close to the assumptions of mainstream phonology. In associationist connectionism, of the kind represented in McClelland and Rumelhart (1986, *et seq*.), and extended to some extent by the models of Hare (1992) and Plunkett and Marchman (1991), the networks themselves have very little *a priori* assumptions, and cognitive processes are built up from data using general-purpose learning procedures. While some assumptions, like the featural basis for unit activation, and the number of nodes, are necessary assumptions to account for the data, the basic idea of these models is that phonological knowledge is built up directly from experience, and very little information is precompiled in it for phonology.

This style of associationism has a problem with implementing variables of the type commonly found in just about every domain of linguistics (see Berent (2013) for extended argumentation). For example, suppose a network is told that AA, DD, ZZ, and GG are grammatical expressions in a language, but that AZ, EG, FE, and SP are not. Suppose then we query the network to see if the novel form EE is grammatical. Most associationist networks, with extensive training, will conclude that EE is not grammatical because EE bears no resemblance to any of the grammatical examples, but bears some similarity to the examples EG and FE. Many phonological systems require this kind of generalization, the most obvious of which is the representation of a geminate, and experimental investigations have shown that humans form this kind of generalization (Berent et al. 2001; Gallagher 2013).

The problem with the XX generalization, where X is some atomic structure in a grammatical expression, is not that connectionist networks cannot represent them. The problem is that they do not induce the pattern from limited data. They cannot generalize the pattern to segments that are not in the training set (Marcus 2001; Pinker 1999; Pinker & Prince 1988; Tupper & Shahriari 2016). In order to handle this kind of "generalization across the board", a network has to have such behavior built into it, such as a mechanism that checks the identity of two segments or that copies material. This has been proposed in some connectionist models (Hare et al. 1995; Shultz 1999); see also Gallagher's (2013) proposal to remedy this problem in Maximum Entropy grammars (Hayes & Wilson 2008).

13.5 Directions for future research

13.5.1 Connectionism and Optimality Theory

In a sense, part of the roots of OT comes from connectionism. The most direct connection is Harmonic Grammar (Legendre et al. 1990), which represents a kind of "half way point" between connectionist networks and OT grammars because of its use of weighted constraints. Digging deeper, though, is the basic idea that grammar can be constituted by a set of constraints. This is a fundamental idea of connectionism because connections serve as constraints on the possible activations of two nodes (Smolensky 1988), and it is also fundamental to OT. Finally, the idea that outcomes are produced by simultaneous evaluation of multiple factors, and assessed for overall harmony, is again central to both models. It is true that most of OT involves symbolic computation, and connectionist networks use numerical computation, but the focus on constraints and parallelism gives the two approaches significant common ground (McCarthy 2002; Smolensky & Legendre 2006).

The examples above that establish parallels between connectionist networks and OT grammars, like the role of the Onset constraint in symbolic and connectionist phonology, are fascinating in their own right, and they bring to the fore the shared principles in the different approaches. These examples are currently few in number, however, and they are based on small fragments of phonological systems. Whether a rich set of parallels exists between the two types of analysis remains to be seen. For example, in the case of Arabic consonant phonology, Alderete et al. (2013) show how the hidden layer nodes "act sort of like" OCP-Place constraints, but the resemblance is not at all exact, and connectionist constraints are in fact laden with context-sensitive effects and exceptions. OT constraints like the OCP are cleaner and do not generally have detailed context sensitivity. The French syllabification and harmony examples present similar segment- and feature-level intricacies that also seem to defy projection to the macro-structure level. On the other hand, other aspects of OT models seem worthy of examination, like the connection between harmonic serialism and recurrence mentioned in section 13.3.6.

13.5.2 Connectionism and exemplar phonology

Another research domain that connectionism can shed some light on is exemplar phonology, or the use of exemplar models of classification and production in phonological analysis (see Wedel (2006) for a review). Connectionist models actually start with very different theoretical assumptions from exemplar models (though some hybrid models do exist, e.g., Kruschke (1992)). As we have seen, connectionist networks involve layers of units linked by connections with varying weights. Connectionist networks implement processes by sending information through a web of nodes and outputting an activation pattern that has an interpretation of some kind. On the other hand, the fundamental unit in exemplar models is the exemplar itself, a detailed memory of some linguistic token. Each exemplar has a location in a space of representations, a label indicating what it is an exemplar of, and a weight indicating how strong the memory of the token is. Linguistic processes include classifying new tokens by their similarity to exemplars already labeled in the representational space, and generating a new token by sampling exemplars according to their weight and reproducing one. Despite these rather different assumptions, there is important common ground between the two models that is useful to understanding how the models work (Bybee & McClelland 2005). First, both are naturally gradient because of their use of continuous variables, i.e., unit activations and connection weights for connectionism, the representational space and exemplar weights for exemplar models. Second, both have an emphasis on learning: information about phonological patterning is built up over stored observations. Third, related to the emphasis on learning, no effort is made to minimize the role of long-term memory. Both models are also task oriented in the sense that knowledge of sound structure is built from normal processes of word learning, which contrasts with more formal models of language learning that de-emphasize the role of memory. Finally, both models have difficulty with making human-like generalization, especially generalization to inputs unlike those in the training data (see Berent's (2013) points on generalization of phonological rules to non-native segments).

Besides these similarities, recent trends in cognitive science are now blurring the lines between exemplar models and connectionist models. A comparatively new development is dynamic field theory, a modeling paradigm built upon connectionist foundations (Erlhagen & Schöner 2002; Johnson et al. 2008). Besides the standard units and connections of connectionism, dynamic field theory introduces neural fields, dense arrays of units interconnected with each other, as well as with other fields and units. Whereas in connectionism each unit has an activation that depends just on time, a field has activation that depends both on time and on the particular location in the field. Neural fields have a combination of excitatory and inhibitory connections that allow the formation of stable peaks of activation, and the location of these peaks of activation can be used to represent the values of continuous variables, such as physical location or color. Furthermore, dynamic fields can be seen as a way of neurally implementing exemplar models. In particular, the activation of a field at a particular location can be used to represent the weight and number of exemplars at a given location in representational space (Jenkins &

Tupper 2016; Tupper 2014). See Spencer et al. (2009) for an overview of dynamic field theory and its rich interaction with connectionism.

13.6 Further reading

This article is intended to provide the rudiments of connectionist networks employed in generative phonology, but there is much more to learn about how these models work. Several chapters from Smolensky and Legendre (2006) provide "need-to-know" background in math and computer science that is extremely useful in understanding how connectionist networks apply to problems in generative linguistics in general, and also psycholinguistics. See also Thomas and McClelland (2008) for a short introduction to the different classes of connectionist networks and how they apply to distinct problems in cognitive science, and McLeod et al. (1998) for an extensive introduction to the discipline, complete with exercises and excellent discussions that flesh out the properties of connectionist networks in classic articles.

For those interested in diving into the details of connectionist phonology beyond the summaries provided here, perhaps the best places to start are the account of Berber syllabification in Legendre et al. (2006), and, for a sequential network, Hare's (1990) model for Hungarian vowel harmony. These two accounts do a fine job of motivating the specific assumptions they make and illustrating how their models produce the outcomes that they do. To learn more about how connectionist networks encode phonological generalizations, Dell et al. (1993) develop a simple recurrent network similar to Hare's model and explain in detail how the model captures the fact that speech errors tend to obey phonological constraints. Goldrick and Daland (2009) and Goldrick (2011) also develop connectionist models for speech errors and explore a number of creative ways in which the structure of speech error patterns can be linked to grammar.

Note

* We are grateful to Sara Finley, S. J. Hannahs, and anonymous reviewers for helpful comments and suggestions. We alone are responsible for any errors that remain.

References

Albright, Adam. 2002. *The identification of bases in morphological paradigms*. Los Angeles: University of California.

Alderete, John, Paul Tupper & Stefan A. Frisch. 2013. Phonological constraint induction in a connectionist network: Learning OCP-place constraints from data. *Language Sciences* 37.52–69.

Anttila, Arto. 2002. Morphological conditioned phonological alternations. *Natural Language and Linguistic Theory* 20.1–42.

Berent, Iris. 2013. *The phonological mind*. Cambridge: Cambridge University Press.

Berent, Iris, Daniel L. Everett & Joseph Shimron. 2001. Do phonological representations specify variables? Evidence from the obligatory contour principle. *Cognitive Psychology* 42.1–60.

Blevins, Juliette. 1995. Syllable in phonological theory. *The handbook of phonological theory*, ed. by J. Goldsmith, 206–44. Cambridge, MA: Blackwell.

Browman, Catherine & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6.201–51.

Browman, Catherine & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49.155–80.

Burzio, Luigi. 2002a. Missing players: Phonology and the past-tense debate. *Lingua* 112.157–99.
Burzio, Luigi. 2002b. Surface-to-surface morphology: When your representations turn into constraints. *Many morphologies*, ed. by P. Boucher, 142–77. Somerville, MA: Cascadilla Press.
Bybee, Joan & James L. McClelland. 2005. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22.381–410.
Clements, George N. & Elizabeth V. Hume. 1995. The internal organization of speech sounds. *The handbook of phonological theory*, ed. by J.A. Goldsmith, 245–306. Cambridge, MA: Blackwell.
Cohn, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7.167–216.
Cohn, Abigail & John J. McCarthy. 1998. Alignment and parallelism in Indonesian phonology. *Working papers of the cornell phonetics laboratory*, 12.53–137. Ithaca, NY: Cornell University.
Corina, David. 1994. The induction of prosodic constraints: Implications for phonological theory and mental representations. *The reality of linguistic rules*, ed. by S.D. Lima, R.L. Corrigan & G.K. Iverson, 115–45. Amsterdam: John Benjamins.
Currie Hall, Kathleen. 2009. A probabilistic model of phonological relationships from contrast to allophony: Ohio State University, Doctoral dissertation.
Dayan, Peter & L. F. Abbott. 2001. *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: The MIT Press.
Dell, François & Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7.105–30.
Dell, François & Mohamed Elmedlaoui. 1988. Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics* 10.1–17.
Dell, François & Mohamed Elmedlaoui. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer.
Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93.283–321.
Dell, Gary S., Cornell Juliano & Anita Govindjee. 1993. Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science* 17.149–95.
Eliasmith, Chris. 2013. *How to build a brain: A neural architecture for biological cognition*. Oxford: Oxford University Press.
Elman, Jeffrey. 1990. Finding structure in time. *Cognitive Science* 14.179–211.
Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge MA: MIT Press.
Erlhagen, Wolfram & Gregor Schöner. 2002. Dynamic field theory of movement preparation. *Psychological Review* 109.545–72.
Fahlman, Scott E. & Christian Lebiere. 1990. *The cascade-correlation learning architecture*. Technical Report CMU-CS-90–100. Pittsburgh: Computer Science Department, Cargnie Mellon University.
Foley, James. 1970. Phonological distinctive features. *Folia Linguistica IV* 1/2.87–92.
Frisch, Stefan A. 1996. Similarity and frequency in phonology: Northwestern University, Doctoral dissertation.
Frisch, Stefan A., Janet Pierrehumbert & Michael B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22.179–228.
Frisch, Stefan A. & Bushra Zawaydeh. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77.91–106.
Gallagher, Jillian. 2013. Learning the identity effect as an artificial language: Bias and generalization. *Phonology* 30.1–43.
Gaskell, M. Gareth, Mary Hare & William Marslen-Wilson. 1995. A connectionist model of phonological representation in speech perception. *Cognitive Science* 19.407–39.
Gasser, Michael & Chan-Do Lee. 1990. Networks that learn about phonological feature persistence. *Connectionist natural language processing*, ed. by N. Sharkey, 349–62. Oxford: Intellect.

Gnanadesikan, Amalia. 1997. Phonology with ternary scales: University of Massachusetts, Amherst, Doctoral dissertation.
Goldrick, Matthew. 2007. Connectionist principles in theories of speech production. *The Oxford handbook of psycholinguistics*, ed. by G. Gaskell, 515–30. Oxford: Oxford University Press.
Goldrick, Matthew. 2011. Linking speech errors and generative phonological theory. *Language and Linguistics Compass* 5.397–412.
Goldrick, Matthew & Robert Daland. 2009. Linking speech errors and phonological grammars: Insights from Harmonic Grammar networks. *Phonology* 26.147–85.
Goldsmith, John. 1992. Local modelling in phonology. *Connectionism: Theory and practice*, ed. by S. Davis, 229–46. Oxford: Oxford University Press.
Goldsmith, John. 1993. Harmonic phonology. *The last phonological rule: Reflections on constraints and derivations*, ed. by J. Goldsmith, 21–60. Chicago: University of Chicago Press.
Goldsmith, John & Gary Larson. 1990. Local modeling and syllabification. *Proceedings of the 26th annual meeting of the Chicago Linguistics Society, part 2*, ed. by K. Deaton, M. Noske & M. Ziolkowski. Chicago: Chicago Linguistics Society.
Hahn, Ulrike & Ramin Charles Nakisa. 2000. German inflection: Single route or dual route? *Cognitive Psychology* 41.313–60.
Halle, Morris & George N. Clements. 1983. *Problem book in phonology: A workbook for introductory courses in linguistics and modern phonology*. Cambridge, MA: MIT Press.
Hansson, Gunnar. 2001. Theoretical and typological issues in consonant harmony: University of California, Berkeley, Doctoral dissertation.
Hare, Mary. 1990. The role of similarity in Hungarian vowel harmony: A connectionist account. *Connectionist natural language processing*, ed. by N. Sharkey, 295–322. Oxford: Intellect.
Hare, Mary. 1992. Phonological representation and processing in connectionist networks: University of California, San Diego, Doctoral dissertation.
Hare, Mary, David Corina & Garrison Cottrell. 1989. *A connectionist perspective on prosodic structure*. Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society, 114–25. Berkeley: UC Berkeley.
Hare, Mary, Jeffrey Elman & Kim G. Daugherty. 1995. Default generalisation in connectionist networks. *Language and Cognitive Processes* 10.601–30.
Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.
Hinton, Geoffrey E. 2016. *Can sensory cortex do backpropagation?* Paper presented at the Information, Inference, and Energy. A Symposium to Celebrate the Work of Professor Sir David MacKay FRS, University of Cambridge.
Idsardi, William. 1992. The computation of prosody: MIT, Doctoral dissertation.
Inkelas, Sharon, Orhan Orgun & Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. *Derivations and constraints in phonology*, ed. by I. Roca, 393–418. Oxford: Oxford University Press.
Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7.217–59.
Jenkins, Gavin & Paul Tupper. 2016. A dynamic neural field model of speech cue compensation. *Proceedings of the 38th annual conference of the Cognitive Science Society*, ed. by A. Papafragou, D. Grodner, D. Mirman & J.C. Trueswell, 508–13. Austin, TX: Cognitive Science Society.
Joanisse, Marc F. 2000. Connectionist phonology: University of Southern California, Doctoral dissertation.
Johnson, Jeffrey S., John P. Spencer & Gregor Schöner. 2008. Moving to higher ground: The dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology* 26. 227–51.
Jordan, Michael I. 1986. *Attractor dynamics and parallelism in a connectionist sequential machine*. Proceedings of the 8th Annual Conference of the Cognitive Science Society, 531–46. Hillsdale, NJ: Lawrence Erlbaum.

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. *A Festschrift for Morris Halle*, ed. by S. Anderson & P. Kiparsky, 93–106. New York: Holt, Rinehart and Winston.

Kruschke, John K. 1992. Alcove: An exemplar-based connectionist model of category learning. *Psychological Review* 99.22–44.

Lakoff, George. 1988. A suggestion for a linguistics with connectionist foundations. *Proceedings of the 1988 connectionist models summer school*, ed. by D. Touretzky, G.E. Hinton & T.J. Sejnowski, 301–14. San Mateo, CA: Morgan Kaufmann.

Lakoff, George. 1993. Cognitive phonology. *The last phonological rule: Reflections on constraints and derivations*, ed. by J. Goldsmith, 117–45. Chicago: University of Chicago Press.

Laks, Bernard. 1995. A connectionist account of French syllabification. *Lingua* 95.51–76.

Larson, Gary. 1992. Automatic learning in a dynamic computational network. *Proceedings of the interdisciplinary workshop on compositionality in cognition and neural networks I*, ed. by D. Andler, E. Bienenstock & B. Laks. Paris: CREA. Ecole Polytechnique.

Lathroum, Amanda. 1989. Feature encoding by neural nets. *Phonology* 6.305–16.

LeCun, Y., J. S. Denker & S. A. Solla. 1990. Optimal brain damage. *Advances in neural information processing systems*, Volume 2, ed. by D. Touretzky, 598–605. San Mateo, CA: Morgan Kaufmann.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Can connectionism contribute to syntax? Harmonic grammar, with an application. *Proceedings of the 26th regional meeting of the Chicago Linguistic Society*, ed. by M. Ziolkowski, M. Noske & K. Deaton, 237–52. Chicago: Chicago Linguistic Society.

Legendre, Géraldine, Antonella Sorace & Paul Smolensky. 2006. The Optimality Theory-Harmonic Grammar connection. *The harmonic mind: From neural computation to optimality theoretic grammar*, ed. by P. Smolensky & G. Legendre, 339–402. Cambridge, MA: The MIT Press.

Marcus, Gary F. 2001. *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: The MIT Press.

Marr, David C. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.

McCarthy, John J. 2000. Harmonic serialism and parallelism. *Proceedings of the North East Linguistics Society 30*, ed. by M. Hirotani, 501–24. Amherst, MA: CLSA Publications.

McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. & Alan Prince. 1986. *Prosodic morphology*. Technical Report No. 32. University of Massachusetts at Amherst and Brandeis University Rutgers University Center for Cognitive Science.

McCarthy, John J. & Alan Prince. 1993. *Prosodic Morphology I: Constraint interaction and satisfaction*. Technical Report No. 3. Rutgers Center for Cognitive Science.

McCarthy, John J. & Alan Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts occasional papers 18, papers in Optimality Theory*, ed. by J. Beckman, S. Urbanczyk & L. Walsh, 249–384. Amherst, MA: Graduate Linguistic Student Association.

McClelland, James L. & Jeffrey Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18.1–86.

McClelland, James L. & David E. Rumelhart. 1985. Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General* 114.159–88.

McClelland, James L. & David E. Rumelhart. 1986. On learning the past tenses of English verbs. *Parallel distributed processing: Explorations in the microstructure of cognition, volume 2: Psychological and biological models*, ed. by J.L. McClelland, D.E. Rumelhart & T.P.R. Group, 216–71. Cambridge, MA: The MIT Press.

McLeod, Peter, Kim Plunkett & Edmund T. Rolls. 1998. *Introduction to connectionist modelling of cognitive processes*. Oxford: Oxford University Press.

McMurray, Bob. 2000. Connectionism for . . . er . . . linguists. *The University of Rochester working papers in the language sciences, 2000(1)*, ed. by K. Crosswhite & J. McDonough, 72–96. Rochester: University of Rochester.

Mitchell, Tom M. 1997. *Machine learning*. Boston, MA: McGaw Hill.

Myers, Scott. 1997. OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15.847–92.

O'Reilly, R.C. 1996. Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation* 8.895–938.

Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. *NELS* 23.367–81.

Pierrehumbert, Janet. 2003. Probabilistic phonology: Discrimation and robustness. *Probability theory in linguistics*, ed. by R. Bod, J. Hay & S. Jannedy, 177–228. Cambridge, MA: The MIT Press.

Pinker, Steven. 1999. *Words and rules*. New York: Harper Collins.

Pinker, Steven & Alan Prince. 1988. On Language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28.73–193.

Plaut, David C. & Christopher T. Kello. 1999. The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. *The emergence of language*, ed. by B. MacWhinney, 381–415. Mahwah, NJ: Lawrence Erlbaum Associates, Ltd.

Plunkett, Kim. 1995. Connectionist approaches to language acquisition. *The handbook of child language*, ed. by P. Fletcher & B. MacWhinney, 36–72. Oxford: Blackwell.

Plunkett, Kim & Virginia Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38.43–102.

Plunkett, Kim & Virginia Marchman. 1993. From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48.21–69.

Plunkett, Kim & Ramin Charles Nakisa. 1997. A connectionist model of the Arabic plural system. *Language and Cognitive Processes* 12.807–36.

Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14.19–100.

Prince, Alan. 1993. *In defense of the number i: Anatomy of a linear dynamic model of linguistic generlizations*. Rutgers University, Center for Cognitive Science.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

Ramsey, William, Stephen Stich & Joseph Garon. 1990. Connectionism, eliminativism and the future of folk psychology. *Connectionism: Debates on folk psychology*, ed. by C. Macdonald & G. Macdonald, 311–38. Cambridge, MA: Basil Blackwell.

Rumelhart, David, Geoffrey E. Hinton & Ronald J. Williams. 1986. Learning internal representations by error propagation. *Parallel distributed processing: Explorations in the microstructure of cognition*. Volumes 1–2, ed. by J.L. McClelland, D. Rumelhard & T.P.R. Group, 318–62. Cambridge: The MIT Press.

Shultz, Thomas R. 1999. *Rule learning by habituation can be simulated in neural networks*. Proceedings of the 21st Annual Conference of the Cognitive Science Society, 665–670.

Smolensky, Paul. 1988. On the proper treatment of connectionism. *The Brain and Behavioral Sciences* 11.1–23.

Smolensky, Paul. 2006. Tensor product representations: Formal foundations. *The harmonic mind: From neural computation to optimality-theoretic grammar*, ed. by P. Smolensky & G. Legendre, 271–344. Cambridge, MA: The MIT Press.

Smolensky, Paul, Matt Goldrick & Donald Mathis. 2014. Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science* 38.1107–38.

Smolensky, Paul & Géraldine Legendre. 2006. Formalizing the principles I: Representation and processing in the mind/brain. *The harmonic mind: From neural computation to optimality theoretic grammar*, ed. by P. Smolensky & G. Legendre, 147–205. Cambridge, MA: The MIT Press.

Spencer, John P., Michael S. C. Thomas & James L. McClelland (eds.) 2009. *Toward a unified theory of development: Connectionism and dynamic field theory re-considered*. New York: Oxford.

Stemberger, Joseph P. 1992. A connectionist view of child phonology: Phonological processing without phonological processes. *Phonological development: Models, research, implications*, ed. by C.A. Ferguson, L. Menn & C. Stoel-Gammon, 165–89. Timonium, MD: York Press.

Stemberger, Joseph P. 2009. Preventing perseveration in language production. *Language and Cognitive Processes* 24.1431–70.

Steriade, Donca. 1982. *Greek prosodies and the nature of syllabification*. Cambridge, MA: MIT dissertation.

Suzuki, Keiichiro. 1998. A typological investigation of dissimilation: University of Arizona, Doctoral dissertation.

Tabor, Whitney, Pyeong W. Cho & Harry Dankowicz. 2013. Birth of an abstraction: A dynamical systems account of the discovery of an Elsewhere Principle in a category learning task. *Cognitive Science* 37.1193–227.

Thomas, Michael S. C. & James L. McClelland. 2008. Connectionist models of cognition. *Cambridge handbook of computational psychology*, ed. by R. Sun, 23–58. Cambridge: Cambridge University Press.

Touretzky, David S. & Xuemei Wang. 1992. *Energy minimization and directionality in phonological theories*. Proceedings of the 14th Annual Conference of the Cognitive Science Society, 248–52.

Touretzky, David S. & Deirdre W. Wheeler. 1990a. *A computational basis for phonology*. Technical Report AIP 113 (ed.) U.o.P. Pittsburgh, PA: Carnegie Mellon University.

Touretzky, David S. & Deirdre W. Wheeler. 1990b. *From syllables to stress: A cognitively plausible model*. Technical Report AIP 117 (ed.) U.o.P. Pittsburgh, PA: Carnegie Mellon University.

Touretzky, David S. & Deirdre W. Wheeler. 1990c. *Two derivations suffice: The role of syllabification in cognitive phonology*. Technical Report AIP 116 (ed.) U.o.P. Pittsburgh, PA: Carnegie Mellon University.

Touretzky, David S. & Deirdre W. Wheeler. 1991. Exploiting syllable structure in a connectionist phonology model. *Advances in neural information processing systems*, ed. by R.P. Lippmann, J.E. Moody & D.S. Touretzky, 612–18. San Mateo, CA: Morgan Kaufmann Publishers.

Tupper, Paul. 2014. *Exemplar dynamics models of the stability of phonological categories*. Proceedings of the 26th Annual Meeting of the Cognitive Science Society, 1628–1633.

Tupper, Paul & Bobak Shahriari. 2016. *Which learning algorithms can generalize identity-based rules to novel inputs?* Proceedings of the 28th Annual Meeting of the Cognitive Science Society, 1289–1294.

Tupper, Paul & Michael Fry. 2012. Sonority and syllabification in a connectionist network: An analysis of BrbrNet. *The sonority controversy*, ed. by S. Parker, 385–409. Berlin: Walter de Gruyter.

Wayment, Adam. 2009. *Assimilation as attraction: Computing distance, similarity, and locality in phonology*. Johns Hopkins University Doctoral dissertation.

Wedel, Andrew B. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23.247–74.

Wheeler, Deirdre & David Touretzky. 1993. A connectionist implementation of cognitive phonology. *The last phonological rule: Reflections on constraints and derivations*, ed. by J. Goldsmith, 146–72. Chicago: University of Chicago Press.

14
Interfaces in connectionist phonology

Joseph Paul Stemberger

14.1 Introduction

Most approaches to phonology within linguistic theory have been focused on classical issues of phonology: how to describe or account for phonological alternations in different morphological or syntactic environments, or how to describe suprasegmental phenomena such as syllable or foot structure. Connectionist approaches to phonology are more often an exception, due to their origin (though see Alderete & Tupper, *this volume*). Connectionist theory was developed initially by cognitive psychologists, with a focus on problems and issues other than phonology per se. (1) Models that included some phonology were developed to address morphology (beginning with English past-tense forms) and naming (a speeded task in which single words are read aloud), and to fit data on language processing in single-word reading or in language production (especially speech errors). While such research sometimes addressed the same issues as phonological theory (e.g. the role of syllable structure), the focus was elsewhere. (2) Most research presupposes that an adequate model must not only account for the relevant data, but also learn from exposure to data, with an active (error-driven) learning process based entirely on domain-general mechanisms. (3) Most models address what humans have in common rather than how they differ. For example, models focus on what is mastered early vs late in acquisition, which at least statistically involves acquisition of one thing before another (such as [p] before [k], or simple before complex onsets). This contrasts with linguistically based research, which most often includes an additional focus of how one system (in one language or in one child) differs from other systems, with the development of typologies; acquisition research often focuses on what the child does *instead of* the adult target (which is highly variable across children). This chapter addresses the interfaces between phonology and connectionist approaches to other things.

The first connectionist models were quasi-symbolic LOCAL or INTERACTIVE ACTIVATION models (e.g. McClelland & Rumelhart, 1981; Dell, 1986; Stemberger, 1985). (A SYMBOL has a straightforward mapping to something in the real world. See below re: quasi-symbolic.) Representations are LOCAL if a concept (semantic, lexical, segmental, featural, etc.) is represented as a SINGLE NODE; thus the word *dog* is represented with a discrete node in the system, as are /d/ and [Coronal]. All nodes have a LEVEL OF ACTIVATION (including an average resting level), and activation SPREADS automatically from one node to another via CONNECTIONS (with WEIGHTS on each connection determining how much activation spreads in a unit amount of time). The structure of the system was simply presupposed; there was no learning component. Models could not start at the beginning and learn the phonology or lexicon of the target language; mature models had to be set up by hand. Such local connectionist models were nonetheless useful, and have continued to be used (e.g. Page, 2000; Magnuson et al., 2007). These models are QUASI-SYMBOLIC because the labels on the nodes correspond to symbols, but processing is not discrete (Fodor & Pylyshyn, 1988): in every use of the system, a stable coalition of elements arises, and it is this GANG that determines the output, rather than just one unit at each level. In essence, the processing of every word involves the entire lexicon, with differences between words reflecting only which lexical unit has the greatest level of activation. Another way to conceive of this nondiscreteness is to consider part of the output the SIGNAL (which the system is intended to output) and part NOISE (which arises spontaneously, is not part of the intended output, and is antagonistic to the signal). In connectionist models, no element is ever at the endpoints and activation is always GRADIENT, meaning that the target elements (most of the signal) are simply the most activated, but many less-activated elements contribute positively to the output (the rest of the signal) or contribute negatively, leading to low-level activation of elements that are incompatible with the target (noise). This is in contrast to competence-based linguistic theories, which presuppose that activation is CATEGORICAL, with some elements at 100% activation and all others at 0% activation, with no noise; such categorical systems can be truly symbolic. Information is not so much STORED as CONSTRUCTED; when it is not being accessed during production or perception/comprehension, information is not in a form that can be used for any purpose whatsoever.

By the late 1980s, the major focus of connectionist modeling shifted to PARALLEL DISTRIBUTED PROCESSING models, which learned the target phenomena from scratch and which contained some nodes that did not map easily onto symbolic elements. Because a concept is represented as a PATTERN OF ACTIVITY across A SET OF NODES, the representations are DISTRIBUTED; there is no single node that corresponds to the word *dog*, but only a particular pattern of activation over a set of units, the same set of units that is used for all words (with each word having its unique characteristic pattern of activation); this is known as SUPERPOSITIONAL MEMORY. The focus on learning distinguishes distributed connectionist research from most research within symbolic models, where it is still acceptable (in 2017) to present results without having to prove that the preferred theoretical model could learn to output the observed patterns on the basis of input.

One major challenge became immediately apparent. Our personal experience with phonology is as a rapidly changing sequence of acoustic (perceptual) or articulatory (production) events, where a given state exists for a short period of time before it is replaced by something else. This is different from our experience with orthographic form in reading, where time is recoded as spatial relations in a two-dimensional space. Linguistic approaches adopt this spatial metaphor: the entire phonological form of a word is simultaneously present, with time encoded along some abstract dimension. The entire time span of a word is available in a usable form, even when not being accessed during production or perception/comprehension. Some connectionist implementations also encode time in an abstract fashion, with different units corresponding to different points in time, so that an entire span of time is available all at once, though, unlike symbolic models, the information is available only when being used. Other implementations encode time in a concrete fashion, as a sequence of different states, only one or two of which exist at a given point in time, as they are being used.

Armed with this basic information, we will first lay out the three types of connectionist models in slightly greater detail, then explore three main areas where they have been applied: morphology, language production, and first language phonological development. I will then explore some predictions of these models. The focus is on production (as are most linguistic approaches to phonology). I will not address connectionist models that have no phonological component, such as models of reading, or even models of naming (e.g. Seidenberg & McClelland, 1989), because there is no discussion of how the phonology works, other than direct ties to the specific task addressed.

Type 1: local connectionist models

All lexical and phonological elements are represented as discrete nodes and are activated simultaneously. For example (Figure 14.1), the node *cat* connects to the nodes /k/, /æ/, and /t/; and /k/ connects to the nodes [Dorsal], [−voice], [−continuant], etc. Note that the features and concepts like syllable structure have simply been borrowed from linguistics, and so are unlikely to directly contribute to our knowledge of units, but that the behavior of the system might in principle tell us something interesting about how phonology behaves in performance.

The base RESTING LEVEL of a node is higher (has more activation) as a function of frequency of occurrence, leading to faster processing and better suppression of competing elements for high-frequency target elements. Co-activated items reinforce each other (SPREADING ACTIVATION), with feed-forward to elements on lower levels and feed-back to elements on higher levels. Because less activation spreads from low-frequency elements, associated elements at other levels also wind up at lower activation levels. There is no MODULARITY (strict separation) between levels, but there is functionally greater separation as the distance between levels increases (Dell & O'Seaghdha, 1991); while phonological features strongly impact on the processing of phonological segments, they have little impact on the processing of semantic elements. Elements on the same level are typically mutually inhibitory, reflecting a COMBAT VIEW OF COMPETITION: e.g., two competing lexical units inhibit each other, but the one that is connected to the larger number of semantic units becomes more activated, and inhibits competitors to low levels of activation. As the nontarget element decreases in activation, it inhibits the target element less (DISINHIBITION), leading to an interactive spiral in which the target element rises to ever greater levels of activation as the competitor falls to ever lower

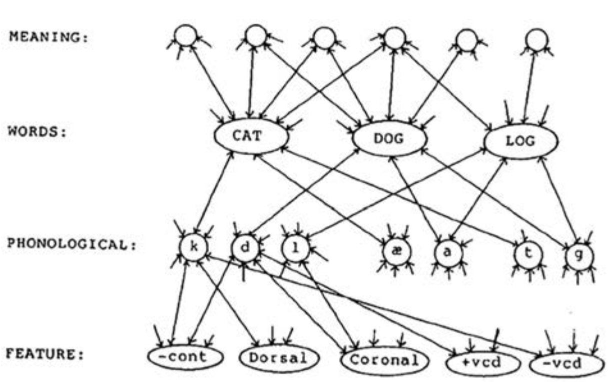


Figure 14.1 Fragment of the system, showing levels and interconnections

levels of activation (the RICH GET RICHER PRINCIPLE). However, activation of nontarget units never falls to zero, because activation continues to flow from elements shared with the target (whether semantic units or phonological features). If nontarget elements are very different, their summed low-level activation pushes the system in all directions and constitutes ineffective noise. If nontarget elements are very similar, they reinforce one output and make it a better competitor; these are consequently known as GANG EFFECTS (McClelland & Rumelhart, 1985). Gang effects can reinforce a target element (and thus speed processing and make errors less likely) or a nontarget competing element (and thus slow processing and make errors more likely). Lastly, if two competing elements are similar in activation level and receive a similar level of input, they can prevent each other from reaching full activation. Usually, this results in blended output at lower levels. In rare instances, activation is insufficient to sufficiently activate elements on lower levels, and we speak of a DELETION error.

Lexical similarity effects (including NEIGHBORHOOD EFFECTS) are ubiquitous in human languages (e.g. Seidenberg & McClelland, 1989; Stemberger, 2004) and arise automatically in local models as one type of gang. Nontarget words are activated via shared units, and the most activated competing words will be those that are semantically or phonologically similar. While the effect in comprehension can be negative (words that are phonologically or orthographically similar interfere with each other and slow processing), in production it is generally positive (with similar words reinforcing each other and making processing more accurate). Seidenberg and McClelland divide these secondarily activated words into friends (which share a particular characteristic with the target word) and enemies (which do not share that characteristic). No connectionist model actually predicts the existence of neighborhoods per se. A phonological neighborhood is defined as the set of words created by making a one-phoneme change to the target word (substituting, deleting, or adding one phoneme); it picks out only the set of most-similar words, is a categorical distinction (words are in the neighborhood or not), and assumes that the three kinds of changes (substitution vs deletion vs addition) have the same impact on lexical similarity, in all parts of the word. In all connectionist models, lexical similarity is gradient (with effects proportional to the number of shared elements), and effects can in principle be different depending on the type of difference and where in the word it is located. This distinguishes research involved with neighborhood density in connectionist approaches (which are more nuanced, and include a distinction between friends and enemies) from linguistic approaches (which generally take all members of a neighborhood as equivalent and as having equivalent effects).

Stemberger (2004) notes that, from a phonological perspective, friends and enemies in neighborhoods differ in terms of coherence. Friends are coherent, because they reinforce the same characteristic, by definition: e.g. in the neighborhood of the word *sick* when we focus on the processing of the /s/, all friends start with /s/ (*sack*, *sip*, . . .). Enemies are generally not coherent, however; words in the neighborhood that do not start with /s/ (*pick*, *tick*, *thick*, *kick*, . . .) each start with a different consonant, and do not reinforce a single competitor; they essentially increase the level of noise in the system, without biasing the system strongly to any particular alternative output. Which words in the neighborhood are friends and which are enemies can be subtle. Stemberger (2004) notes that for the target word *sick*, words in the neighborhood that start with an alternative consonant nonetheless do reinforce the fact that the ONSET contains a single consonant; while they are enemies of the consonant /s/, they are friends of the simple ONSET.

Gangs give the lie to the characterization of local models as "local", as opposed to "distributed". All lexical items are processed whenever any lexical item is processed, and together all lexical items influence the output. Functionally, the representation of any particular lexical item is distributed across all the lexical items in the system.

There are two basic ways to encode time in these models. First, the system can split into multiple SLOTS, with elements duplicated in each slot, allowing for the access of multiple items simultaneously. Second, the system's output can change over time, activating first one element then another. Dell (1986) mixes these two aspects. He presupposes that all the words of a sentence are activated by the semantics at the same time, but that one word is selected as the "current node", and its activation is boosted. Once processing is complete enough, the word is selected and copied to a frame, and the activation of its component elements is set to zero. A new "current node" is then selected. Processing thus changes with time, and real time is an integral part of the access of all units. If the current word node has several syllables, one syllable is set to the "current node" and its activation is boosted. Within a syllable, however, all segments are accessed independently in different slots, presumably in a CCCVVCCC frame. There are thus different units for the "same" phoneme in different positions in the syllable, with different feature units, and the prediction is that elements in these different slots are processed independently and do not spontaneously interfere with each other in any way (whether positive or negative). Note, however, that a variant model might carry the "current node" mechanism down to the segment level, which would allow different segments in the same syllable to interact spontaneously.

Type 2A: Distributed nonrecurrent networks

A typical nonrecurrent net, which was used for very few models, is shown in Figure 14.2.

Rumelhart and McClelland (1986) developed the first nonrecurrent distributed model to address morphology. The input was a distributed representation of the base word's pronunciation (e.g. /wɑk/ *walk*, and /goʊ/ *go*). Because the entire word was output in one pass, time had to be represented in an abstract fashion, and WICKELFEATURES (context-sensitive

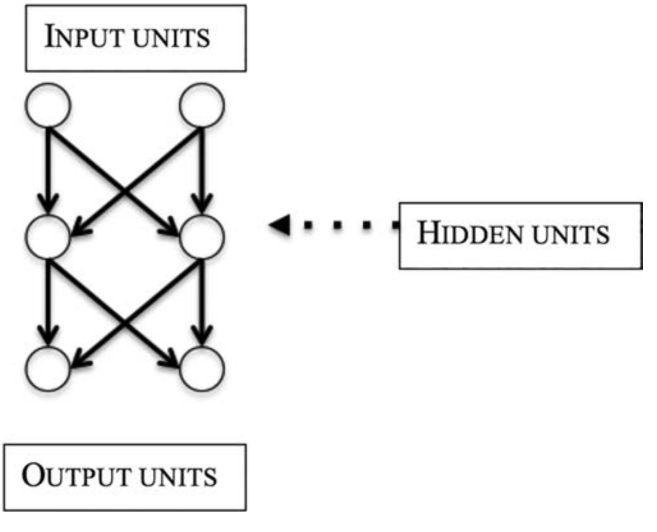


Figure 14.2 Basic nonrecurrent net, with two units in each layer; for additional units, add connections to all units in other layers. Each connection has an independent weight

allophones; see section 14.2.2) were arbitrarily chosen. These were mapped directly onto an output representation of the pronunciation of the past-tense form (/wɑkt/, /wɛnt/); all input nodes connected to all output nodes. Both regular and irregular forms were stored in the same set of units and connections. Activation flow was entirely unidirectional (feed-forward), flowing from input to output and never the reverse. Like local models, the basic architecture of the system had to be hand-tailored. The advantage over local models was that there was an algorithm (BACK-PROPAGATION) for learning content; the exact output pronunciation of each past-tense form could be learned on the basis of input. The algorithm is one variation of ERROR-DRIVEN LEARNING: when the system recognizes that the output is in error, it alters the weights between units to make it less likely that an error will happen next time. Because making large changes leads to instability of output, changes were small, so observable developments occurred slowly over a long period of time, during which there is variability between correct and incorrect outputs that gradually approaches almost total accuracy. Error-driven learning entails a component that can recognize accurate production vs errors, which Rumelhart and McClelland called "the teacher". This was misunderstood by e.g. Pinker and Prince (1988) as requiring overt correction from adults, which they note rarely occurs and is rarely effective. However, the intent was an internal mechanism that monitors output, recognizes error, and provides information for learning; in the language-production literature, this is known as a MONITOR (e.g. Baars et al., 1975; Levelt, 1989; Hartsuiker, 2006). There is evidence that adult speakers have two monitors: one that operates in a pre-articulatory fashion to detect and prevent errors before they are uttered, and one that operates on the basis of sensory feedback (to identify errors after they occur).

Frequency effects arise in this system automatically. High-frequency words have more learning trials than low-frequency, so come to be more accurately processed with less error. High-frequency patterns (such as regular *-ed*) are subject to more learning trials, leading to lower error rates than with low-frequency (irregular past-tense) patterns. High-frequency patterns are also more likely to overgeneralize (over-regularization errors such as *breaked* for 'broke'), though groups of phonologically similar irregular forms also form a locally frequent pattern that can overgeneralize in that restricted part of phonological space (over-irregularization).

Type 2B: distributed recurrent networks

More recent connectionist models are most often recurrent in nature: they contain loops that allow multiple passes through the system, with each pass corresponding to a phoneme (in models that generate the phonological output of words). The basic structure of such a system is illustrated in Figure 14.3 with the model of Dell et al. (1993), which was designed to address phonological processing; the model takes meaning as input and gives phonological features as output.

Each word is a pattern of activity across semantic feature nodes (which were actually random, because no database of semantic representations for lexical items are available) that is mapped onto a pattern of activity across phonological feature nodes, and these two layers are mediated by a layer of hidden units that maps the input onto the output. This model has two context layers to keep track of previous outputs. Note that the input and output units are symbolic, but that the hidden units are nonsymbolic.

In principle, the simplest recurrent system can be made up of two levels (input and output), plus the mechanism needed for the implementation of time. In practice, networks work better if there is a self-organizing layer in between: the hidden units. All implemented

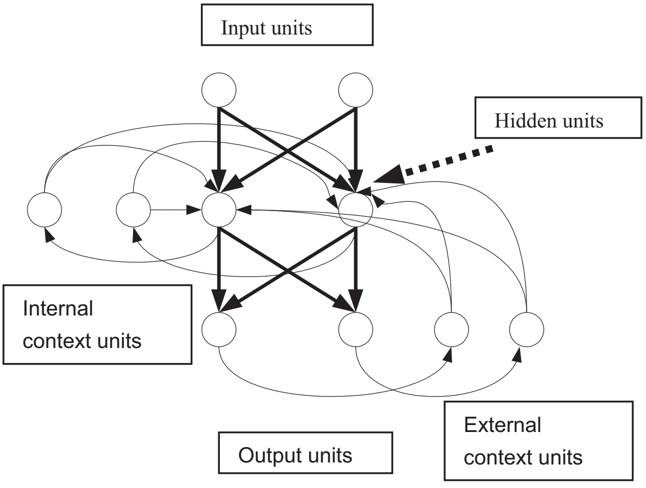


Figure 14.3 Recurrent net, with both internal and external context units added

connectionist models have been hard-wired for this, with a specific number of hidden units. While the input and output units are arguably symbolic, corresponding to something observable in the world (such as semantic features in the input, and articulatory units in the output), the hidden units learn the best connection weights to connect the input and output, and so do not correspond expressly to any observable in the world. An optimal learning system would not have to know in advance how many hidden units are needed, or even that one layer of hidden units is needed (rather than two). In principle, there can be multiple layers of hidden units, which organize themselves into useful intermediaries. One could imagine word-like, syllable-like, and segment-like units intermediary between meaning and articulation.

Activation flow oscillates between feed-forward and feed-back, with much more coarse-grained time intervals than in local models (e.g. Elman, 1990). The system outputs a single segment at a time. The pattern of activity in the output layer is stored in an EXTERNAL CONTEXT layer; Elman used only this type of context layer. The pattern of activity in the hidden units is saved in an INTERNAL CONTEXT layer; Corina (1994) used only this type of context layer. How many context layers should exist, and where they should recur to (as input to the hidden layer or as input to the output units), has not been systematically explored. In the Dell et al. model, the first pass through the system outputs the first segment of the word, though Jordan (1986) notes that each pass could be smaller (e.g. one millisecond). Both context layers are then input into the hidden units on the second pass through the system, along with the *same* semantic pattern that was used during the first pass. That semantic pattern, in combination with context information about what was just produced, leads to the output of the second segment of the word on the second pass; without the context layers, the output would be the same as on the first pass. The pattern of activation in the hidden units and output units in the second pass are stored in the context units and input through the hidden units on the third pass, resulting in the third segment of the word. This continues until the system returns a null element that corresponds to the word boundary. E.g., for the word *cat*, given that the system has just output /k/, the semantic pattern for {cat} leads to the output of /æ/; given that it has just output /æ/, the semantic pattern leads to the output of /t/; and the fourth pass returns a null element, signaling that the word has reached its end.

The context units enable the system to extract generalizations across a series of passes, whereby one type of output state narrows the range of possible output states in the next pass. For example, after a pause, virtually any output state is possible; but suppose that /p/ was the actual output. On the second pass, only [+sonorant] speech sounds are possible in English, specifically /l, r, j/ and vowels; if the actual output on the second pass is /l/, then the next pass must output a vowel. The context units thus extract probabilistic phonotactic information on sequences of speech sounds.

Neighborhood effects (or rather, lexical similarity effects) arise naturally on the basis of the processing of similar words. The units that lead to such interactions are those between the hidden units and the output units, as well as both sets of context units.

An optimal recurrent system is driven by meaning-to-sound mappings. Time is encoded directly in an analog fashion, and on the pass that outputs phoneme #3, only phoneme #2 exists fully (in the external context units), a trace of phoneme #1 remains (in the internal context units), and all later phonemes in the word are only potential and do not exist in detail.

14.2 Morphology

There are many papers that implement models of morphology, all within the distributed variant of connectionism; they address accuracy in learning and generalization of morphological patterns. The ability to handle phonology was not a goal; they do not address the learning of the basic phonology of the language, nor accuracy on things irrelevant to morphology, such as word-initial consonant clusters, velars, or words with consonants of different places of articulation. Shortcuts were generally made, with no attempt to argue for the structure of the system. There was far more hand-tailoring (and symbolic representations) than is generally acknowledged. In this section, I focus on what the models had to say about phonology (and phonology–morphology interactions). Phonological effects on morphology that derive from error-driven learning will be addressed in section 14.5.

14.2.1 Local models

There has been some (non-implemented) discussion of morphology in local connectionist models. Stemberger (1985, 1994) noted that local models are compatible with a rule-based approach to morphology (i.e. there is a lexical unit such as *-ed* that associates past-tense semantic units with phonological units), but could in principle generalize morphological patterns even if e.g. past-tense forms were stored as wholes. The issue of whether rules exist or not is beyond the scope of this paper on phonology, but one point is perhaps relevant. In the case of semantically similar lexical elements such as target *walk* and competitor *run*, the target has to suppress the competitor cleanly enough so that the phonological units of the competitor have relatively minor effects on output: [wɑk] is output rather than e.g. [wɑn] (a blend of the two words, with [wɑ] from *walk* and /n/ from *run*). However, if all inflected forms of a word are stored separately (e.g. *walk*, *walked*, *walks*, *walking*), interaction during processing can be very different. Because e.g. *walk* and *walks* share [wɑk], processing needs to be accurate enough that the end of the word matches the target inflectional category, but further suppression of the nontarget inflected form is often unnecessary. Stemberger suggests that different inflected forms of the same word can form a gang that makes access of the shared phonological material easier. Similarly, while *run* and *ran* must compete for the vowel to be output, they can cooperate on access of the /ɹ/ and /n/.

Stemberger (1994) discusses generalization of the allomorphy of e.g. past-tense -*ed* (/əd/ after /t, d/; /t/ after other voiceless obstruents; /d/ after other voiced segments). Even if known past-tense forms are stored as separate lexical units, gangs will form during processing that generalize across them. For example, if there are 1,000 nontarget past-tense lexical units, most end in -*ed*, with the frequency of allomorphy being /d/ > /t/ > /əd/. Hundreds of past-tense forms each contribute a small amount of activation to /d/. The /d/ unit sums this activation, and the result is that /d/ gets more activation than any phoneme within the target word. However, when /d/ is unlikely because it creates a consonant cluster that is impossible in English (as in **walk-d*), it attains a lesser degree of activation, and the second most frequent past-tense pattern, /t/, wins. Spreading activation activates uninflected forms with similar sequences (such as *act*), creating a tendency for inflected forms to match general output limitations in the lexicon (such as legal /ks/ but illegal */kz/ and */gs/). If neither /t/ nor /d/ is phonologically possible (as in **need-d* or **need-t*), then the least frequent regular pattern, /əd/, wins. Output reflects general patterns in the lexicon, both for uninflected words and inflected words, with predicted interactions across all types of words.

Regular morphological patterns are generally associated with hundreds of lexical items, and so have enough base support for ready generalization, leading even to over-regularization of irregular past-tense forms (e.g. *breaked* instead of *broke*); failure is greatest for low-frequency irregulars, whose phonological information is least-well-encoded. Irregular patterns generally have fewer exemplars (with no more than 30 in irregular past-tense patterns in English), and so none gains enough activation to be a good general competitor. However, a phonologically based gang can form for irregulars, if a number of irregulars end in a particular sequence of vowel+consonants that does not end in -*ed*. E.g., families of irregular past-tense forms that end in *ank* and *unk* (etc.) can generalize to new forms (at a low frequency), even leading to over-irregularization (e.g. *grun* instead of *grinned*; cf. *spun*, *run*).

14.2.2 Distributed models

Distributed models of morphology have attracted the greatest attention, because of their learning component. Rumelhart and McClelland (1986) set the scene with a model that was intended solely to learn past-tense forms, with a goal of showing that the system could generalize to new words, and over-regularize irregulars. The system took the phonology of base forms as input, and gave as output the phonology of past-tense forms. One can reasonably ask whether it is legitimate to posit such a network. On the basis of semantic information, why should the phonology of uninflected base forms first be accessed, which are then altered into past-tense forms by a secondary network? What leads the system to create a separate network for some lexical outputs but not others? A particular phonological representation was assumed, with no justification: wickelfeatures, context-sensitive features such as $_{[-\text{continuant}]}[+\text{continuant}]_{[-\text{continuant}]}$, which would be found in a continuant segment (such as a vowel or fricative) surrounded by noncontinuants (such as stops or nasals). Pinker and Prince (1988) heavily criticized the use of wickelfeatures, and they were rarely used again; the main drawback is that they hinder generalization across the same segment in different environments. But immediately, models (e.g. MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991; Daugherty & Seidenberg, 1992) followed Rumelhart and McClelland in (a) not justifying the existence of a dedicated subnetwork for inflected forms, (b) not justifying the particular representation of phonology that is used, and (c) not addressing how well the system learns phonological aspects of the target languages (such as complex onsets like /pl/ and phonemes like /k/). Hare and Elman (1992) focused on the learning of the morphology without including the phonology of the base in the output. Their model of plural forms in Old English had outputs just for the suffixes and for the change of a vowel to a front vowel (as in *foot/feet*), eliminating all aspects of the pronunciation that were irrelevant to plural morphology. While the output units were a subset of the units that would be found in a full model, and should behave exactly like the same subset of units in a full model, it provides no information about the general phonological behavior of the model.

Rumelhart and McClelland (1986) did, however, provide some information about the accuracy of phonological outputs in their model, noting that it produced some odd-looking forms such as [mɛmbəld] as the past tense of the target word *mail* /meɪl/. Pinker and Prince (1988) noted that the system seemed to have problems in general with the "vowel-l" region of phonological space, and described such outputs as "grossly bizarre". Pinker (1984) and Pinker and Ullman (2002: 458) contrast these outputs with the outputs of "humans" and "people". However, what is known about phonological development in children suggests that outputs such as *membled* are expected to occur in the speech of some children.

Plunkett and Juola (1999) used a slightly expanded task, with the phonology of base forms as input and with a layer of hidden units, and the network learned to do both noun plurals and verb past-tense forms (suggesting equivalence to the final lexical stratum in Lexical Phonology; Kiparsky, 1982). The entire pronunciation was output all-at-once, with time recoded in an abstract fashion. All inputs were monosyllabic, with a maximally CCCV-VCCC form. For each of the allowed eight positions, there were 16 binary phonetic features. Segments were "left-justified" in each of the three areas (onset–nucleus–coda); e.g. *walk* as /w-0–0=ɑ-0=k-0–0/. No reasons were given for choosing left-justification. The use of separate sets of output units for different positions has the detrimental consequence of isolating each position from all others during processing: there is no sense in which a unit such as [+voiced] in position-1 is the same as the unit [+voiced] in position-2, and a feature in position-1 could not spontaneously lead to assimilation or dissimilation of features in position-2. Plunkett and Juola do not address the accuracy of the phonological output in general, but do focus on two aspects relevant to English morphology: the presence of a schwa in words such as *bushes* and *needed*, and the voicing of the plural and past-tense suffixes (*-t* and *-s* after voiceless segments, *-d* and *-z* after voiced). The presence/absence of schwa was accurate in both trained words and untrained words. They imply that voicing was accurate in trained words, but had a high (9.2%) error rate in untrained words (and do not address the level of accuracy in over-regularization errors such as *breaked*, *knifes*, etc.). This error rate is higher than expected in the speech of even the youngest children who can differentiate the voicing of final obstruents, suggesting that the model does not match human phonological behavior.

Plunkett and Nakisa (1997) addressed noun plurals in Arabic, examining the productivity of a suffixing plural (which accounts for only about 24% of noun stems, but generalizes most readily to new words) vs the "broken plural" (a set of 31–70 patterns in which the consonants are the same as in the singular form, but the vowel patterns are different from the singular). An input template VCVCVCVCVCVCV, with each consonant of the stem assigned to C elements in order, allows the input consonants to be mapped directly on to the output consonants, and to allow the vowels to change. No reasons are given for different justification rules in Arabic vs English, but it should be noted that this justification resembles that of McCarthy (1981) for morphological templates in Arabic. Had the system had to change the order between Cs and Vs more directly, it would have been a more difficult problem for learning. Plunkett and Nakisa reported that the system generalizes morphologically as humans have been claimed to generalize, but no details of the phonological behavior of the system were given.

In relation to this last point, it should be noted that these nonrecurrent models do not predict that general phonological frequency in the language will affect morphological patterns. Because these models have a special dedicated network, the sole purpose of which is to create inflected forms, they can pick up on statistical properties of the relevant inflected forms only. If, in contrast, the network produces all words (uninflected and inflected and derived), the system would extract the statistics for all words. Stemberger and Middleton (2003) argued that general vowel frequency in the lexicon as a whole accounts better for the processing of irregular verbs in English than does vowel frequency restricted to just past-tense forms. Stemberger (2007) showed that overtensing errors (e.g. *I didn't broke it*, with inflected *broke* in place of nonfinite *break*, in a past-tense context) are far more common with regular verbs if they create a rime that occurs in uninflected words (e.g. *kissed*, cf. *mist*) than if they create a rime found only in inflected forms (e.g. *buzzed*, *liked*). Such evidence suggests that models with subnetworks that generate only inflected forms do not match the behavior of actual humans.

Bernhardt and Stemberger (1998) present an instance of interaction between a child's possible word-final consonant clusters and the double expression of the past-tense *-ed*. This child was able to output word-final sequences ending in /t/ or /d/ preceded by a vowel (including diphthongs), or a sonorant (/ɹ/, /l/, or nasal). The child produced past-tense forms with a single /t/ or /d/ when they fit exactly this form. With all other within-morpheme consonant sequences ending in /t/ or /d/, the child simplified by deleting one of the consonants (generally the /t/ or /d/); e.g. *act* [ʔækʰ]). But the child could produce many consonant sequences in word-medial position. Past-tense forms that ended in other consonants, which should have led to impossible word-final clusters, were altered into acceptable word-medial clusters by doubling the *-ed*; e.g. *picked* [pʰɪktəd]. In Dell et al.'s (1993) model, this makes sense. In a target word such as *picked*, the third pass would output [k], and the fourth [t], but this is not a possible word-final sequence in the child's speech. Since the input is still the same, the system has the option of continuing to a fifth pass, outputting the *-ed* a second time (since past-tense forms frequently end in [təd] or [dəd]), and continues until [pʰɪktəd] is output; this shifts an impossible word-final cluster into medial position, where it is a possible output. We see the interaction of knowledge about past-tense forms interacting with the general phonological ability of the child. Again, this suggests that any model with a subnetwork restricted to past-tense forms cannot capture actual human behavior. A single network is required to output all words, of all morphological compositions.

Currently, no implemented connectionist model of morphology has addressed general phonological properties of a language, nor been part of a system that produces all the words of a language. Insofar as they address the morphophonemic alternations that draw most attention in phonological theory, they have met with limited success. Unimplemented local models have been explored that do this to some extent, but address only basic phonological processing.

14.3 Language production

There is a literature that addresses phonological processing in language production, aimed at (a) how speakers produce phonological plans for target words, (b) the phonological speech errors that arise during that process, and (c) the way that the system malfunctions after brain damage. Dell (1986; Dell et al., 1993) has worked with implemented local and distributed models, and Stemberger (e.g. 1985, 1991a) has explored such models without actually implementing them. Phonological effects in language production that derive from error-driven learning will be addressed in section 14.5.

The goal of these models was to fit data from phonological production, in particular data from phonological speech errors. Researchers analyzed speech error data (from corpora of errors in natural speech, or using experimental error-induction tasks such as SLIP). One interpretation of such data is that virtually anything is possible, but certain types of errors are of quite high frequency, while others are quite rare. The goal of the models was to be able to derive all known types of errors and match their frequencies, thus providing one explanation of why/how such errors occur, and also to make predictions about factors that should affect speech errors. I will address major issues in turn, but leave several that result from error-driven learning for section 14.5.

14.3.1 Basics

Phonological errors occur on a trial when a target unit is inaccessible (due to natural variability around the mean resting level) or when there is a strong competitor that is activated for independent reasons, or a combination of these reasons. When a target element is inaccessible, it may simply achieve such a low activation level that it disappears (a NONCONTEXTUAL DELETION ERROR). Alternatively, a competitor may be activated enough to outcompete the target, leading to a NONCONTEXTUAL SUBSTITUTION ERROR. If there is interference from an independently activated element, a CONTEXTUAL SUBSTITUTION ERROR results; the statistically most common independent source of activation is because the element appears in a nearby word. A nontarget element may appear, either for unknown reasons (a NONCONTEXTUAL ADDITION ERROR) or because it is independently activated, generally because it appears in a nearby word (a CONTEXTUAL ADDITION ERROR). Lastly, a target element may be deleted due to interference from a nearby word where that element does not appear (e.g. in target /pl/ the /l/ is deleted, because a nearby word has target /p/); this is viewed as interference at the syllable-structure level, leading to loss of a slot for the /l/. Syllables may also be lost (generally in a noncontextual fashion).

14.3.2 Noise suppression: covert contrasts in errors made by adults

There is some indication that substitution errors may leave low-amplitude traces of the original target features (e.g. Pouplier & Goldstein, 2005). Errors most commonly occur via interference between two words, with two consonants (or two vowels) undergoing far greater competition than usual. There is every reason to expect that this more than usually intense competition should result in a lower than usual activation of the winning nontarget phoneme, with less than usual suppression of the losing target features (Stemberger, 1992b). The result is a phoneme with adequate articulation of the nontarget feature, but with some low-level articulation of the target feature. Due to this residual noise, e.g. the substituting phoneme [k] is not expected to be phonetically identical to true target [k]'s.

14.3.3 Frequency effects

All other things being equal, high-frequency elements are processed more accurately than low-frequency elements, due to their high resting activation level (local models) or greater number of learning trials (distributed models). Stemberger and MacWhinney (1986); Dell (1990); and Petrič and Stemberger (2014) report that low-frequency words are more likely to undergo phonological speech errors than high-frequency words. Countless studies have shown that low-frequency irregular inflected forms are more likely to be regularized than high-frequency forms (e.g. Bybee & Slobin, 1982; Marcus et al., 1992; Marchman, 1997; Petrič & Stemberger, 2014); and that the low-frequency irregular patterns are themselves more subject to error that the high-frequency regular pattern.

In local connectionist models, the greater activation of high-frequency units not only makes processing more accurate on the unit itself, but cascades down to all associated elements, increasing their accuracy as well. Thus, not only are high-frequency words more accurately accessed, but so are their component phonemes (Stemberger & MacWhinney, 1986; Dell, 1990).

14.3.4 Segmental similarity effects

Given competition between two elements, the intruding element is more likely to win if it is of high activation, and one source of activation is from components (such as features) that are shared with the target element. All studies show that, other things being equal, a target consonant or vowel is more likely to be replaced by another segment that shares many features. For example, when processing a /b/, substitution errors are quite likely when there is a nearby /p/ (differing only in laryngeal features), /v, m/ (differing by just manner features), or /d, g/ (differing by just place features); less likely near /f/ (differing in both laryngeal and manner features), /z, ð, ʒ, dʒ, n, ŋ, l, ɹ, w, j/ (differing by both place and manner features), or /t, k/ (differing by both laryngeal and place features); and even less likely near /s, θ, ʃ, tʃ, h/ (differing in laryngeal, manner, and place features).

14.3.5 Segment vs feature errors

Errors are more likely to involve whole segments than individual features. This derives from the nature of feedback in local models. Every feature spreads activation back to segments that are connected to it, spreading activation to nontarget segments that share features with the target, but also providing the most feedback to the target segment itself. If one feature fails, the target segment gets less activation from features, and the other features of that segment become more likely to fail as well. Segment units thus create a tendency towards "all-or-nothing" processing, where all features are successfully accessed or none are. Feature errors do occur, but at a much lower rate than whole-segment errors, both in natural speech and in the SLIP task (Stemberger, 1991b).

14.3.6 Identity effects

Linguistic theory presupposes that there is a mechanism that can recognize segments as a unit, and that constraints on repetition can e.g. cause dissimilation of one of the identical segments. Because local connectionist models contain a unit corresponding to e.g. the lexical item *walk* or the phoneme /b/, they provide a locus for IDENTITY EFFECTS. A particular unit has just been used, meaning that it reached a high level of activation. If there is a period of time during which the unit's activation has not yet dropped back to its original resting level, it will be easier to access a second time, a phenomenon known as IDENTITY PRIMING. Alternatively, after being used, the activation level might be reset to a very low level to prevent accidental re-use (perseveration), a phenomenon known as a REFRACTORY PERIOD, during which the unit would be more difficult to access a second time. These effects are distinct from effects of similarity.

Stemberger (2009) argues that, while the existence of a similar non-identical segment in a nearby word increases the likelihood that that source segment will appear in the target word as a substitution, the existence of an identical segment in a nearby word makes it less likely that that segment will appear. Phrased another way, a /b/ in a nearby word makes it more likely that a /b/ in the target word will undergo a substitution error. Based on similarity effects and spreading activation, one might erroneously have expected that a nearby /b/ would decrease the rate of errors on a target /b/. There appear to be two separate effects (similarity and identity), with different impacts on processing, as local models predict could be the case. Stemberger shows that the magnitude of the repetition effect is identical on Onset consonants, coda consonants, and vowels, suggesting a single mechanism. Interestingly, there is no effect of the repetition of features in vowels or coda consonants, suggesting that the locus for the repetition effect involves segment units, which he suggests is tied to the probability of nonsystematic repetition across words, a factor that also affects the directionality of errors.

14.3.7 Directionality of errors

To implement interference on the target word from previous vs upcoming words, Dell (1986) used two parameters: A (anticipation) vs P (perseveration). Increasing A leads to more anticipation errors, while increasing P leads to more perseverations. Exchanges (where, in the sequence AB, phoneme A appears early and phoneme B appears late) occurred through an interaction of the A and P parameters. Dell reported that it was impossible to set the parameters so that phonological exchanges are the most common type of error. This at first appeared to be a problem, because Shattuck-Hufnagel and Klatt (1979) emphasized the role of exchanges (on analogy with the stress on word exchanges by Garrett, 1975). However, Stemberger (1989) noted that completed exchanges were the least common type of error, behind completed anticipations and perseverations. The most common type of error is actually incomplete anticipations, where the speaker stops and self-corrects between the error and the word that is the source of the error; if most of those errors are exchanges, then exchanges are the most common type of error; but if most of the incompletes are anticipations, exchanges are the least frequent type of error. Stemberger suggests that it is plausible that the incompletes are made up of anticipations and exchanges in the same proportion as completed errors, if the self-correction is based primarily on the first error, and not on any upcoming error. If incompletes are treated as anticipations vs exchanges in the same proportion as completed errors, the proportion of anticipations vs exchanges is similar for adults and young children for phonological errors, with exchanges being the least common type of error. (In contrast, word sequencing errors are primarily exchange errors for speakers at all ages.) Stemberger (2007) has suggested that exchanges are the result of a mechanism designed to prevent perseverations, a mechanism whose strength is related to how often elements are repeated. Words are repeated within sentences at a low rate; phonemes are repeated at a somewhat higher rate (ca. 9% for word-initial consonants, and ca. 15% for vowels and word-final consonants); while phonological features are repeated at very high rates (with the probability of repetition greater than for some features). Observed errors in natural speech show the rate of exchange errors inversely proportional to the probability of accidental repetition: high for words, low for phonemes, extremely low for features.

Dressler (1979) notes that one striking characteristic of speech errors is that the two interacting consonants can ignore similar intervening consonants. This is because there is interference between two words at a basic level, not directly derivable from the linear order of the phonemes. He notes that this is quite different from most long-distance assimilations, which often cannot skip over intervening material (cf. also Gafos, 1995).

14.3.8 Neighborhood effects

As discussed above, neighborhoods (or rather, similar words in proportion to their similarity to the target word) are expected to have an effect, with (coherent) enemies increasing error rates and friends decreasing error rates. Vitevitch (2002); Stemberger (2004); and Petrič and Stemberger (2014) show that this is the case. Stemberger notes that when the relevant error involves a substitution (e.g. *sick* ➔ **shick*), having more friends decreases error rates, while the number of enemies (which do not reinforce any particular consonant) has no effect. When the error involves an addition (e.g. *sick* ➔ **stick*), words such as *thick* and *pick* are actually friends that make up a coherent gang that reinforces the target simple Onset, and the more there are, the lower the rate of addition errors.

14.3.9 The repeated phoneme effect

Local models do not just predict that similar competitor words in the lexicon will be activated. Nearby words that are similar to the target word also receive spreading activation from the target word. Similarities such as shared vowels increase the activation level of the nearby word, which spreads activation to the competing phonemes, which increase the likelihood of errors. Thus, given a target word such as *pen*, and nearby words such as *best* or *bust*, the models predict that the /b/ of *best* is more likely to lead to a substitution error (*pen* ➔ **ben*) than the /b/ of *bust*. See e.g. Dell (1986); Stemberger (1990); and Petrič and Stemberger (2014) for empirical verification of this predicted effect. In most instances, this seems to have a similar effect as the constraint contiguity in Optimality Theory: an intruding phoneme is more likely to replace the target phoneme if the next phoneme is the same, so that contiguity is not violated.

14.3.10 The parallel syllable structure constraint

Dell (1986) notes that every study has found that segments statistically tend to interact with segments from a parallel part of the syllable: onsets with onsets, nuclei with nuclei, codas with codas. The issue of slots in syllable structure is an interesting one, and has been adopted as a basis for the frame-content theory of MacNeilage (2015, a recent review). In Dell's model, this is built in as a categorical effect, due to having parallel and non-interacting slots within the syllable. Unfortunately, there are exceptions. Stemberger (1985) notes that syllabic /ɹ̩/ acts as the source for errors involving addition of /ɹ/ to create Onset clusters, substitution for vowel targets, and addition after a vowel to create sequences such as /ɑɹ/. There may also be a small percentage of cross-syllable-position errors. Any constraint must be gradient rather than categorical. A distributed model such as Dell et al. (1993) may also predict this effect, since material from comparable passes in interacting words are most likely to interfere.

14.3.11 Syllable structure errors (addition/deletion)

Errors occur in which segments are added or deleted. Stemberger (1990; Stemberger & Treiman, 1986) reported that, when /t/ and /ɹ/ compete (e.g. due to priming in a SLIP experiment in a word-pair such as *tamp roll*), expected substitution errors such as *ramp roll* occur, but addition errors such as *tramp roll* are almost as common. The intuition is that competition between two phonemes can be eliminated by creating another slot to accommodate both phonemes. But they also note that such an error is only common when the interfering consonant would appear in second position (as in *tramp roll* above) and not in first position (as in *tamp troll*); indeed, they observed a minority of errors like *tramp toll*, where a C2 slot was created for the error /ɹ/, but no C1 could be added for the error /t/. This is possible in the system of Dell (1986), because feedback from a highly activated error phoneme can activate an alternative syllable structure with a complex Onset, with the observation that a target singleton C is in the C1 slot and rarely is shifted to a C2 slot, so it is easier to add a C after a single C than before it. Dell et al. (1993) give up the concept of slots and syllable structure, and present a recurrent system that outputs one phoneme at a time. It is not easy to see how competition from a consonant in another word would lead to an addition error; and if it did, it isn't clear why the interfering consonant can be added as C2 but much less often as C1. In *tamp roll*, if /t/ is output on the first pass in the second word, what prevents /ɹ/ from being output on the second?

Stemberger (1990; Stemberger & Treiman, 1986) also reported that consonants could be lost from clusters under some circumstances, and noted that for adults the lost consonant tended to be C2, e.g. for *tramp toll* the error *tamp toll* is quite common, but for *tramp roll* the error *ramp roll* in uncommon; and for *stamp sole* the error *samp sole* is quite common, while for *stamp toll* the error *tamp toll* is uncommon. Again, this suggests stability of C1 and instability of C2. And again, it's unclear how the distributed model of Dell et al. (1993) would deal with such phenomena.

14.3.12 Lexical blend errors

An interesting prediction is made concerning contextual synonyms. Consider the following sentence:

(1) It has a very nice *flaste*. 'flavor' & 'taste'

In this sentence, both alternative words, *flavor* and *taste*, mean roughly the same thing, and either could be used to express the meaning. Rather than settle on either word, both competing words are highly activated, and continue to compete at the phonological level. All reports agree on two characteristics (e.g. MacKay, 1973; Stemberger, 1985; Laubstein, 1999). (1) Most errors begin as one word and end as the other, and do not switch back and forth. This presumably reflects a tendency for sequences of segments to cohere in some way; the OT constraint contiguity would have this effect. Other than the contiguity violation at the switch point between the two competing words, this constraint would be unviolated. (2) At a far greater than chance level, there is a phoneme at the switch point that is the same in both words, as in the example above. Segment-to-segment contiguity is not violated even here. In local models, the shared phoneme effect is due to spreading activation involving the phonemes of the two words, but there is no direct explanation for the fact that the speaker does not switch between the words multiple times. In Dell et al.'s distributed model, once a switch has been made, the context units will provide information that will keep the output coming from the same word. If the two competing words have the same vowel, the external context units do not provide any information to differentiate the two words; the internal context units will be most compatible with the word that was output first, but are less distinct than if there were no phoneme in common.

14.4 Phonological development

Linguistic theories have been designed to allow the description of the phonological form of words in adult pronunciation, both basic characteristics of all words (such as the particular inventory of speech sounds in a particular language, and how they can be combined into syllables and words) and alternations (wherein a word is associated with some differences in the pronunciation of some specific part of the word in some specific environment). But the theories potentially make predictions about the acquisition of phonology by young children. A major issue is whether the system may have particular tendencies, and whether those tendencies correspond to what children do spontaneously. A further question is whether these natural tendencies might be extended to account for phonological change in historical linguistics. Phonological effects in child phonology that derive from error-driven learning will be addressed in section 14.5. Stemberger (1992b) presents an unimplemented local approach and discusses developmental phenomena, and Berg and Schade (2000) present an implementation designed specifically to deal with consonant harmony. Plaut and Kello (1999) present a distributed model designed to get at phonological development, and make reference to some developmental phenomena (but not in any detail). This section explores what sorts of phenomena might be expected to arise spontaneously.

14.4.1 Basics

Early in development, from a connectionist perspective, some adult targets are inaccessible in the child's system, while others are accessible. Connectionist models should show frequency effects: high-frequency things should be learned earlier than low-frequency things (because there are more learning trials, or because resting activation levels are higher). They are expected to show effects of similarity across lexical items (e.g. neighborhood effects), because all lexical items are produced using the same set of units and so naturally influence each other (positively or negatively, depending on the characteristic in question). If an output is inaccessible, it might simply be deleted. If the level of activation is high enough, however, a substitution will result, especially replacement by a high-frequency (often early-learned) segment, due to spreading activation between target features and competing phonemes. Note that an output might generally be accessible, but be replaced by a particular highly accessible competitor in particular circumstances. Note that accessible units might be barely accessible, in which case they may tend not to occur as substitutions for inaccessible targets (even when expected). In the local model of Dell (1986), one would also expect errors in the syllable structure, especially errors where (lower-frequency) complex onsets, nuclei, or codas are simplified to (high-frequency) simple units. In distributed models, there is natural competition between the different segments in the word, and it is predicted that phonological elements might appear in the wrong place in the word, possibly as apparent assimilations (anticipatory or perseveratory).

14.4.2 Noise suppression: covert contrasts in child phonology

In a child substitution error (e.g. *see* [t^hiː]), the system has failed to access the low-frequency feature [−continuant], outputting high-frequency [−continuant] instead. But note that the feature that is output is not supported by learned lexical connections; it is accessed purely on the basis of generalization across other words (and frequency information extracted by the context units). It is thus predictable that the activation of the mismatching segment [t^h]-for-/s/ will have a lower activation level than a target [t^h]-for-/t/ in a word such as *tea*, which, in addition to the general information that leads to [t^h] in the word *see*, is supported by lexical connections. In a connectionist system, an error element can never exceed the activation of a target element, and will usually have less activation. Because an error [t^h] has less activation than a target [t^h] (with less suppression of competing features such as the target [+continuant]), this difference can lead to weaker articulation that is less [−continuant]-like than a target [t^h]. Subtle phonetic differences between errors vs targets (covert contrasts) have been reported for phonological development (e.g. Macken & Barton, 1980; Tyler et al., 1993; Scobbie et al., 2000; Munson et al., 2010). Covert contrasts are expected to occur in connectionist systems.

14.4.3 No modularity between phonology and phonetics

Local models presuppose interaction between adjacent levels of the system, with both feed-forward and feed-back. If the level of phonological features feeds forward to a phonetic/motor system, the motor system will feed back to the phonetics. Phonological units will thus be influenced by motor units: a highly accessible motor unit spreads activation back to the phonological units and reinforces them; an inaccessible motor unit spreads no activation back, leading to lower activation of the target feature and a higher probability of error. In this way, phonological elements can be sensitive to motor difficulty. Note that there are often multiple explanations for why particular outputs are difficult. Voiceless fricatives are often mastered before voiced fricatives. Voiceless fricatives are simpler to produce, because they require high oral airflow to produce fricative noise with no secondary source of sound. Voiced fricatives are more complex to produce, because they require high oral airflow to produce fricative noise, but relatively lower pressure above the larynx than below, in order to produce voicing. But the voiceless fricative /s/ in English is also more frequent than the voiced fricative /z/. A connectionist model does not allow reduction to phonetic difficulty (which here is actually aerodynamic difficulty rather than motor difficulty); processing factors cannot be ignored.

14.4.4 Babbling: initial training of the system

While the majority use of the system is meaningful speech, there are exceptions. For example, it can be used for repeating back words whose meaning is unknown or unpredictable, including new names (a natural task) or novel words (as happens in experimental tasks). In the earliest output, beginning at about 0;6, infants babble, initially a purely motor action. It has been shown that there is continuity between babbled "utterances" and the first meaningful words (e.g. Locke, 1983; Vihman, 1996): the first words often bear a close phonetic resemblance to babbling. This implies that the output units are the same for babbling and early words (e.g. MacNeilage, 2015; Guenther, 1995). Plaut and Kello (1999) present a distributed model where training of phonological output units begins with learning the acoustic consequences of babbling.

14.4.5 Quirky outputs (membled)

Pinker and Prince (1988) protested outputs like *membled* as the past-tense form of *mail*, but young children often pronounce words in ways that are quite different from adults, and it is not an obvious problem that the phonological output of a developmental model is less than adult-like. When examining odd phonology for past-tense forms, we can separate out the accuracy of the phonological form from the accuracy of the morphology. The form [mɛmbəld] ends in [d], and is of the form that is expected in words with a similar output structure, such as *mumbled*. The system can be viewed as getting the morphology right, but not the phonology. Young children can and do produce inflected forms where the morphology seems accurate but the phonology is odd. Bernhardt and Stemberger (1998) reported the following plural forms from a child at 3;2:

(2) *orange* /ɔɹəndʒ/ [ʔaːzət]
oranges /ɔɹəndʒəz/ [ʔaːzəts] ~ [ʔaːzəz]

Note that the two plural variants fit the general pattern for English (/s/ after /t/, /z/ after a vowel). The variant [ʔaːzəts] is reasonable given the child's rendition of the singular form, while the variant [ʔaːzəz] is a closer approximation of the adult form. The word *orange* is phonologically unique in English, but on the basis of most other words in the child's speech with similar sounds, the predicted pronunciations were singular [ʔoʊwənz] and plural [ʔoʊwənzəz] (which emerged two months later). The fact that the Rumelhart and McClelland model output the same sort of quirky phonological forms that children do cannot be considered a bad thing.

Recent linguistic approaches to phonological development using Optimality Theory maintain that all complex patterns in phonological development can be broken down into components that are shared to different degrees with other children, creating typologies of possible systems (e.g. Bernhardt & Stemberger, 1998; Pater & Barlow, 2003; Barlow & Gierut, 1999). In evaluating whether a given hypothetical pronunciation is possible, we ask whether the constraints that lead to phenomena that have been observed in actual children can be combined in a hypothetical child in a way that will lead to that hypothetical pronunciation. As regards a word like *mailed* /meɪld/, children learning North American English generally do not have [l] in codas until quite late; Smit (1993) reports only 50% correct usage in the 3;6–5;0 age group. Children often "vocalize" coda /l/ after a diphthong to a central or back vowel, and *mailed* would commonly be output with two syllables as [meɪ.oʊd]. Bernhardt and Stemberger (1998) note that children sometimes require syllables to have onsets; this most commonly involves the extension of the glide portion of a diphthong (here /eɪ/) to the Onset of the second syllable ([meɪjoʊd]), but may also involve a copy of a consonant elsewhere in the word: [meɪmoʊd] (cf. *piano* [pʰɪnæːno]). This is enough to establish the behavior of the model as reasonable, but I note that other attested phenomena could get the pronunciation as far as [mɛmbʌd]. There is just one aspect of [mɛmbəld] that is very unlikely: it probably outputs a pronunciation that is *less* like the adult form than [mɛmbəld].

Rumelhart and McClelland's model thus deviated from the target forms in ways reminiscent of real children, but was outputting such forms in the final state of the system, and so failed to attain a final system like those of adult English speakers. But a model should behave like adult humans only if it has had the same level of experience with past-tense forms. We can estimate the approximate age of the model, given the amount of training that it had. The model had been trained 190–200 times on 336 regular verbs, for a total of about 67,000 separate attempts. How old would a human be with that level of training? Wagner (1985) estimates that a child uses about 14,000 word tokens per day. An analysis of the data from Adam, Eve, and Sarah (Brown, 1973) and Abe (Kuczaj, 1977) suggests that about 0.6% of young children's word tokens are regular past-tense forms (2,684 of the 462,182 words spoken by the children). If the first regular past-tense form is attempted at 1;10, the child will reach 67,000 tokens in about 797 days, at about 4;0. The approximate age of the model is 4;0. A child's phonology is still far from adult-like at 4;0, but really odd pronunciations tend to be at a bit younger age; the child above who produced [ʔaːzəts] for *oranges* did so at 3;2. At worst, the model is slightly less mature than actual children.

14.4.6 Less similar substitutions

If a feature such as [+continuant] is inaccessible in a fricative such as /f/, leading to a stop, it might be expected that place of articulation might be relatively unaffected: a labial stop would be output, most likely shifted to a bilabial to correspond to other system outputs ([pʰ]). While this is the case for the majority of children, it is not at all uncommon for a dental [tʰ] to result (e.g. Bernhardt & Stemberger, 1998). There are two explanations for this in connectionist systems. (1) In local models, features feedback to reinforce the activation of the target phoneme (Dell, 1985). If a target feature is inaccessible, there is less feedback, leading to lower activation of the target phoneme, which in turn gives less activation to its other features. So the loss of [+continuant] in /f/ leads to a lower activation level for [Labial], which may also be lost. Activation of [Labial] is lowered even further if the target secondary place feature [+labiodental] is also lost (because stops are [−labiodental]). In essence, segment units create a tendency towards all-or-none access; if one feature fails, the others may also fail. Given low activation of [Labial], the high-frequency competitor [Coronal,+anterior] may win instead, leading to a substitution that is less similar to the target segment than expected. (2) There is a second mechanism in both local and distributed models. Suppose that [pʰ] is accessible, but barely; it is a weak output, with very little extra margin. Given that activation is lower in errors than in targets, it may be that [pʰ] is simply not strong enough to be accessed without support from lexical connections. Given that [pʰ] is inaccessible, the system outputs highly accessible [tʰ].

14.4.7 Syllable structure

In phonological development, complex onsets initially are absent, and fewer consonants are produced. The typical result is to retain the phoneme of lower sonority for most children (as predicted by the Sonority Sequencing Principle, but also by the frame/content theory of MacNeilage, 2015): *spot* and *plot* as [pʰat]. Less commonly the two consonants are coalesced (e.g. [fat], combining the place and voicing and obstruence of the /p/ plus the [+continuant] feature of the /s/ or /l/), or both consonants are deleted ([at]). In a local model (e.g. Dell, 1986), there can be an error where only a C1 slot is accessed, which can accommodate only one consonant. This creates competition between the two consonants, which is generally resolved if one wins, but the winner is determined more by high-frequency features than by original position in the word; but the winner wins at the phoneme level. Alternatively, the competing features can lead to secondary activation of a third phoneme unit (such as [f]), and the features that win are the low-frequency nondefault features that have especially strong lexical connections). The fact that coalescence prefers low-frequency nondefault features, while if C1 or C2 wins the preferred features are the high-frequency default features [−continuant] and [−sonorant], suggests the role of both phoneme units and feature units. If competition is particularly strong, both competitors can inhibit each other to such a strong degree that no consonant gains enough activation, and so both consonants are deleted. In the distributed model of Dell et al. (1993), it might be possible to expect that high-frequency features such as [−continuant] and [−sonorant] would win on the first pass, but it isn't clear why e.g. reduction of /sp/ to [p] is more common than coalescence to [f], nor is it clear how both consonants could be deleted.

14.4.8 Consonant harmony

Vihman (1978) explored characteristics of noncontiguous consonant assimilation in child phonology. A commonly reported instance is the word *dog*, where the target is an anterior coronal and the second is a velar, where the child outputs both as a velar: [gag]. This is not a problem with the phoneme /d/, since this is generally mastered far earlier, and indeed was generally present even in the word *dog* at an earlier age. Bernhardt and Stemberger (1998) treat it as a problem with coordinating two different places of articulation in the same word. In the distributed model of Dell et al. (1993), consonant harmony would arise in the following way. On the first pass through the system, both [Coronal] and [Dorsal] are "latent" in the output, since the input leads to both [Coronal] (the target for the first pass) and [Dorsal] (the target for the third pass). In [gag] (VELAR HARMONY), the feature [Dorsal] is erroneously output on the first pass. If the system is not prone to exchange errors (and few children show such METATHESIS phenomena), then [Dorsal] will also be output on the third pass (where it belongs). It has been shown that the target of place harmony is most likely to be high-frequency [Coronal] rather than lower-frequency [Labial] or [Dorsal], and that harmony is more often anticipatory (right-to-left) than perseveratory (left-to-right). We return to such statistical tendencies below.

While it is uncommon, Stoel-Gammon (1996) reported a child with Velar Harmony which was restricted to words in which a back vowel intervened between the two consonants; an intervening front vowel blocked the harmony, as in *duck* [g̊ʌk] but *stick* [tʰɪk]. Because the harmony occurs on the first pass, it is not yet known what the vowel will be; the reason for the harmony is competition of the information that belongs to different passes, bypassing any intervening segments. The intuition is that the spread of [Dorsal,+back] is facilitated in *duck* because the vowel can also participate in the assimilation, while this is impossible in *stick* because the vowel is [−back] and cannot participate in the spread of [+back]; but it is entirely unclear how this could be relevant in a distributed recurrent model. Stemberger (1993) reports an instance of Labiodental Harmony, where a labiodental [f] causes a nearby (preceding or following) bilabial /m/ to become labiodental [ɱ]: *smell mice* [fɱɛʊ ɱaɪs] and *small home* [fɱaʊ hoʊɱ]. However, the assimilation occurred only if the intervening segments were [−consonantal] and so lacked consonantal place features (whether vowels or glottals such as /h/), and was blocked by intervening consonants with oral place features: *smelled mice* [fɱɛʊd maɪs] and *small comb* [fɱaʊ kʰoʊm]. Again, a distributed model has no mechanism to deal with such effects.

14.4.9 Interference in long words

There is a further implication for especially long words. Long and complex words have so many passes, with so many latent phonemes, that it is quite likely that competition will be especially high, and so children should be particularly challenged. James (2006) and Mason (2015) show that this is the case.

14.5 Detailed exploration of one characteristic and its consequences: error-driven learning in connectionist systems

As the speaker produces language, a monitor assesses the accuracy of the output. Presumably, the reference is initially (in very young children) the perceived form, which would involve an EXTERNAL MONITOR (assessing the acoustic consequences of articulation). With increasing experience with articulation, an INTERNAL MONITOR becomes possible, assessing the internal plans (lexical, morphological, phonological, phonetic) and possibly responding to sensory feedback from the vocal tract. Traditionally, error is assessed in a categorical fashion: was the input correct or not? If the monitor decides that no error has occurred, the system is not changed. In a complex system that is functioning correctly, changing the system can potentially induce errors that had not occurred previously. If the monitor decides that an error has occurred, connection weights between input and output are altered slightly, to make the erroneous output less likely on the next trial. Changes to connection weights are small, because large changes can induce errors that had not occurred previously, and make the system unstable over time. Over thousands of trials, the small weight changes add up, and the output converges on the reference. However, back-propagation assesses error in a gradient fashion: how far the output was from the reference. Due to noise in processing, the output is never perfect. With the gradient way of assessing error, there is always error, and there are weight changes after every trial, leading to lifelong learning in an attempt to attain the unachievable perfection.

14.5.1 Perfect vs good enough

One can adopt an intermediate criterion for accuracy: once accuracy has reached a certain level, and the degree of error is sufficiently small, the system takes the output as “good enough”. Since noise can never be eliminated, there is little gained by pursuing noise suppression beyond a certain level. If altering weights in learning consumes cognitive resources, this would be expected. This raises the following question: how good is good enough?

14.5.2 Intelligibility

One answer is that close to 100% intelligibility (in typical speech contexts) is good enough. Noise that has minimal impact on intelligibility is tolerated. Most noise is nonsystematic, involving slightly lower activation for target units and slightly higher activation for nontarget units. Such variation can perhaps safely be tolerated.

14.5.3 Systematic noise: covert contrasts (adult)

In adult language, when a morpheme is associated with multiple phonological outputs, there is always activation of the variants associated with nontarget inflected (or derived) forms, especially of high-frequency variants. Consider word-final devoicing of voiced obstruents in e.g. Slovenian:

(3)		nom.sg.	gen.sg.
	led- ‘ice’	[let]	[leda]
	polet- ‘flight’	[pɔlet]	[pɔleta]

In processing of *polet*, all inflected forms have a voiceless final obstruent, there is strong activation of [−voiced] and no strong source of activation for [+voiced]. In the processing of *led*, there is lexical activation of [+voiced], which will maintain a much higher level of activation even if the system requires [−voiced] to be output. The effect at the phonetic level is a less extreme degree of voicelessness, with some characteristics that are slightly more voiced-like. Even if it is possible to suppress this systematic noise completely (and it probably isn't), the suppression of the noise would have to go well beyond any consideration of intelligibility (where, indeed, any small voiced-like characteristics that can be perceived actually facilitate correct identification of the target word). Similarly, in an English phrase such as *good boy* [gʊb̚bɔɪ], the output unit [Coronal] of the /d/ of *good* will be activated by the word *good*, even as the unit [Labial] is assimilated from the /b/ of *boy*. It is unclear if the [Coronal] unit can ever be suppressed fully (given continuous lexical activation from the *good* unit), and such suppression makes no functional difference. One can still accurately describe the events as "devoicing" and "labial assimilation"; the slight phonetic differences constitute a tolerable level of noise, with subcategorical differences.

Linguistic models (and those in speech science) lead one to treat everything that someone produces as intentional, as part of the plan, as signal (rather than noise). Allowance is occasionally made for some characteristic to arise within the vocal tract due to physics and not as part of the motor plan created in the brain; e.g. the fact that VOT for [p^h] is longer before high vowels than before low vowels due to the effects of differences in cavity size on the speed of pressure equalization. Connectionist models allow for systematic differences to arise also during the planning phase, due to the way that lexical items map onto speech sound units.

14.5.4 Multiple sources of activation: underspecification

Information is stored in more than one location and combined during the construction of the output: two lexical locations (input-to-hidden-unit connections; hidden-unit-to-output connections) and two context locations (internal-context-to-hidden-unit connections; external-context-to-hidden-unit connections). In principle, the information that leads to accurate output is split between the lexical and context connections, but is not split uniformly for different target segments. The context units allow the system to predict characteristics of the next pass. As a simple example, in a language with only open syllables and no complex onsets, where every word is of the form CVCV(. . .), after a C is output, it is predictable that a V will follow (and vice versa). Because the context units provide support for a vowel output, error is lower and less learning takes place in the lexical pathway: [consonantal] is distributed between lexical and context pathways, and so is represented less strongly in the connection weights in the lexical pathway. This is akin to phonological underspecification in theoretical phonology (where predictable phonological features do not appear in lexical representations), except that it is a gradient version (with weak rather than absent representations). This extends to high-frequency features as well. If the consonant feature [Coronal] is more frequent than the competing units [Labial] and [Dorsal] (as it probably is in all human languages), context units will more strongly support [Coronal], because it is the most likely feature to follow. This leads to increased accuracy (lower error) and less learning, so that [Coronal] is distributed across lexical and context pathways and so is more weakly represented in the lexical pathways. In contrast, the less-frequent [Labial] unit is not as strongly supported by context and in fact must deal with interference from the context units (which provide more activation for the competitor [Coronal]); error is greater, more learning occurs, and [Labial] is more strongly concentrated in the lexical pathway.

Consequently, the highest-frequency competitor has a weaker lexical representation than lower-frequency competitors, but the highest-frequency competitor is nonetheless processed more accurately, because of strong support from the context pathways. Stemberger (1991a, 1991b; Stemberger & Stoel-Gammon, 1991; Petrič & Stemberger, 2014) discuss underspecification in these terms. Jakobson (1941/1968) notes that unmarked features are generally of higher frequency than marked features, though some phonologists maintain that defaults can be of low frequency within a language (de Lacy, 2006). Bybee (2001) objects to underspecification, on the grounds that a mechanism would be needed to remove predictable elements from lexical entries, and notes that a usage-based system of storage has no such mechanism; but usage-based systems need a mechanism that will store all information strongly. Connectionist systems do not require a mechanism to remove predictable elements, nor do they require a mechanism to store predictable elements strongly in a lexical fashion. If the goal of learning is accuracy of output, it does not matter where or how information is stored, only that the information is highly accessible.

14.5.5 Consequences for competition-intensive situations

Because high-frequency unmarked default features have weaker lexical representations, they do not compete well with lower-frequency marked nondefault features (with strong lexical representations) in situations where there is strong support for the nondefault feature. In grammars, in instances of assimilation, nondefault features have strong lexical support, and so can outcompete default features; it is predicted that assimilation should more often involve spread of nondefault features to replace default features rather than the reverse (Archangeli, 1988; Archangeli & Pulleyblank, 1994; Paradis & Prunet, 1991). Similar asymmetries are predicted for speech errors, errors involving irregular verbs in English, and consonant harmony in child phonological development.

14.5.5.1 Phonological processing

Similar asymmetries are predicted for speech errors where features whose activation is split across lexical and context connections are expected to undergo errors whereby they are replaced by features with strong lexical connections. For example, target anterior coronal /t/ is expected to be replaced by nearby labial /p/ or velar /k/ more than the reverse; but errors between labial /p/ or velar /k/ are expected to show more errors on the phoneme of lower frequency (/p/ in this case). Stemberger (1991a) showed that this holds true of speech errors in spontaneous speech and in the SLIP task. Petrič and Stemberger (2014) show that it holds true of dental vs palato-alveolar phonemes in Slovenian. The results, as predicted by connectionist models, is a mixture of frequency effects and apparent anti-frequency effects, as a function of how concentrated activation is in the weights of lexical connections. Mid front vowels, despite being of very high frequency in English, are predicted to lose the competition with other vowels.

14.5.5.2 Phonological competition in irregular verbs

Stemberger (1993) reasoned that competition between the base and past-tense vowels of irregular forms might be resolved phonologically, in favor of the vowel with stronger lexical connections. Stemberger (1992a) had first shown that such an effect was present in speech errors, and on the basis of that adult (experimental) data, made predictions about which irregular past-tense forms would be most likely to undergo over-regularization and base-form errors. For example, the following prediction was made:

(4) HIGH rate of over-regularization: *fell* → *falled*
LOW rate of over-regularization: *broke* → *breaked*

Stemberger, in a CHILDES-based study, found that this phonological factor had as strong an effect as lexical frequency and membership in hypersimilar families. Marchman (1997) showed this in an experimental study. Stemberger and Middleton (2003) reported the same results experimentally in adult speech.

Stemberger and Middleton extended this to a different type of error: overtensing errors such as the following:

(5) LOW rate of over-regularization: *didn't fall* → *didn't fell*
HIGH rate of over-regularization: *didn't break* → *didn't broke*

This looks like a syntactic error, where the main verb occurs with an auxiliary and so should have an uninflected form, but where the child double-marks tense on both the auxiliary and the main verb. If there should be a bias towards vowels with stronger lexical connections, it is predicted that the statistics of overtensing should be the opposite of over-regularization: if the vowel of the base form tends to be retained, it will show up in overregularizations (high error rate) and with auxiliaries (low error rate); if the vowel of the base form tends to lose to the vowel of the past-tense form, it will not show up in overregularizations (low error rate) but will show up with auxiliaries (high error rate). This was demonstrated experimentally for adults, and for children by Stemberger (2007). Both phenomena taken together show that phonological processing takes place at the same point in time as irregular forms are constructed, and that the results are affected by biases in phonological processing.

14.5.5.3 Consonant harmony in phonological development

Stemberger and Stoel-Gammon (1991; Stoel-Gammon & Stemberger, 1994) demonstrate that consonant harmony in child phonology is more likely to affect the high-frequency default categories of alveolar place of articulation and stop manner of articulation, replaced by lower-frequency labial and velar places of articulation and fricative and nasal manners of articulation; consonant harmony involving two lower-frequency places or manners of articulation more often do not show an asymmetry of this sort. This is as expected, since features whose activation is spread between lexical and context connections are losing the competition to features whose activation is concentrated in lexical connections. Berg and Schade (2000) present a local model of consonant harmony in which connection weights are stronger on [Labial] and [Dorsal] than on [Coronal], and derive the observed asymmetries.

For distributed models, consider Dell et al. (1993). All features are potential on every pass, since the same semantic input is present for every pass; a learner must learn to output a particular feature only on a particular pass. In the first pass for words such as *dog* and *top*, [Coronal] should be output, but note that [Labial] and [Dorsal] have strong lexical connections. It is thus possible that [Labial] and [Dorsal] will win on the first pass (*gog*, *pop*), thus creating Labial or Velar Harmony. On the third pass, because of strong lexical connections, we expect the [Labial] or [Dorsal] feature to also be output where it belongs; unless of course there is a mechanism that prevents possible cases of "perseveration", or even specifically makes sure that a given feature in a word is output only once, in which metathesis (*god*, *pot*) will result. Consonant harmony is much more common in child speech than metathesis (also known as feature migration), so we can conclude that most children have systems that do not lead to metathesis.

As a qualification, Vihman (1978) notes that consonant harmony is twice as likely to involve anticipation (a feature appears early) than perseveration (a feature appears late). In "classic" child systems, the words *dog* and *top* undergo consonant harmony, but the words *cut* and *pot* do not. This implies that once a feature is output in its target location, it is much less likely to interfere with later segments. Stemberger (2013) investigates word-internal nonsystematic speech errors in child speech, and notes the same statistics: segments or features are much more likely to be anticipated than perseverated (contrasting with between-word nonsystematic speech errors, where perseverations are far more common than anticipations).

Error-driven learning makes an error less likely as learning progresses. However, when an error becomes impossible, it is not necessarily the case that the correct output will appear. The output may reflect a different attractor state (especially one that already exists as a possible output); in some cases, this will involve u-shaped learning. Consider the following developmental progression:

(6) *house* [haʊh] > [haʊtʰ] > [haʊs]

Initially, faced with the impossibility of a fricative, the child matches [+continuant] at the cost of losing oral place of articulation ([Coronal,+anterior]) and obstruence. When this error is suppressed, place and obstruence reassert themselves, but fricatives are still impossible, resulting in a stop. Developmentally, we see the output [+continuant] being replaced by [−continuant], to re-emerge at a later state when fricative manner is mastered. An instance of u-shaped learning for the feature [continuant] (but straight improvement for the features [−sonorant,Coronal,+anterior]).

14.5.6 Similarity between sounds

In processing, segments that share many features interact more than segments that share few features, giving a measure of similarity. It is reasonable that if a target segment is inaccessible, the system will output a similar phoneme instead. But the system must correct such substitutions, which makes them interact less, and leads to changes that could be described as making the two segments less similar. Consider the following developmental sequence:

(7) *coin* [tʰaɪn] > [tʰeɪn] > [tʰɔɪn]

The child initially substitutes an unrounded central vowel for the target back vowel, at a time when the front diphthong [eɪ] was not present in the child's system. When [eɪ] appeared, the child began substituting [eɪ] for /ɔɪ/, reflecting the fact that the vowels are of a similar height, though they differ in backness. But most adults find [eɪ] less similar to /ɔɪ/ than [aɪ] is. To suppress one error, the system is now behaving as if the degree of similarity between vowels has changed. Contrast this with a path in which [aɪ] was maintained until [ɔɪ] was mastered. The dynamics of these two systems, once [ɔɪ] has been mastered, will not necessarily be the same. One might expect that the similarity relations between [ɔɪ] and the vowels [aɪ] and [eɪ] will be different, as a function of the child's developmental pathway. This suggests that similarity functions may differ across adults, reflecting differences in learning in early child phonology.

14.6 Conclusions

We have explored characteristics of connectionist models (local, distributed nonrecurrent, and distributed recurrent), with reference to phonology, which have been developed for reasons distinct from those of phonological theory. These models (or explorations) have focused on morphology, on language production (often to account for the statistical properties of speech errors), and on phonological development in young children. Some models have focused on phonological issues that arise in language production, phonological development, or morphology, but for many phonology has been an afterthought. Some of the characteristics stressed by connectionism have more recently been the focus of work that is better known to phonologists, through the work of laboratory phonologists. In this final section, I will overtly bridge the gap between these connectionist models and theoretical linguistic approaches to phonology. One important thing to consider is whether the characteristics discussed above were deliberately built in to connectionist models, or emerged from other assumptions. The built-in characteristics are most often borrowed from phonological theory directly, and may not be very informative for phonological theory. The characteristics that emerge have more interesting things to say to phonological theory.

1. Frequency effects are intrinsic at all levels (lexical items, phonemes, and features), plus combinations such as diphones.
2. Similarity between lexical items is intrinsic. However, unlike neighborhoods, similarity is a gradient measure.
3. Similarity between segments affects the likelihood of interaction, even for accidental interference in speech errors. The use of a constraint such as AGREE, motivated by the effect of shared features, is unnecessary.
4. Segmental similarity is likely to be perturbed by error-driven learning, such that the contribution of each feature to similarity will vary from speaker to speaker (and probably from language to language). A universal similarity scale is unlikely.
5. The “assimilations” of phonological processing can skip over intervening material that one might consider “relevant”. Models based on such behavior allow such skipping. This suggests that what is going on in such processing models is quite different from phonological alternations of the traditional sort, which usually disallow such skipping. In addition, not everything that can be viewed as an assimilation is due to spreading.
6. There are effects of segmental identity, suggesting that distinguishing between whole segments (with the OCP and with the constraint IDENT) and features (with the OCP and with the constraint MAX) is reasonable.
7. The effects of contiguity arise in connectionist models (in blend errors and with the repeated phoneme effect), suggesting that a constraint such as CONTIGUITY is reasonable.
8. Syllable structure effects may arise automatically in distributed models due to extraction of statistics from segment sequences by the external and internal context units. Local models build in syllable structure position effects by hand (which isn’t very interesting), but then predict that the number of slots in syllable structure can be in error (cluster simplification in child phonology; contextual addition and deletion in speech errors). This resembles syllable-structure-related changes in phonological theory, as well as in reduplication.
9. Phonological effects on morphological form arise intrinsically in processing, much as would be expected on the basis of allomorphy effects in adult languages.
10. Due to differences in activation due to lexical vs general sources, changes in pronunciations (whether adult errors or child phonology phenomena) are not expected to be phonetically identical to similar underlying pronunciations. Covert contrasts are expected because processing is gradient, not categorical. Terms such as “substitution” and “deletion” in principle are not contradicted by the existence of covert contrasts.

11 Error-driven learning leads to underspecification (in a gradient version), and with it a whole host of phenomena in phonological processing and in phonological development.

The phonological behavior of connectionist models designed to address unlearned phenomena (speech errors, developmental phenomena) and incidental phonology (in models of morphology) does not correspond exactly to theoretical phonological approaches. They do, however, provide an interesting perspective and an interesting comparison.

References

Alderete, J., & P. Tupper. (This volume). Connectionist approaches to generative phonology.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*, 183–207.

Archangeli, D., & D. Pulleyblank. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Baars, B. J., M. T. Motley, & D. G. MacKay. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382–391.

Barlow, J., & J. Gierut. (1999). Optimality theory in phonological acquisition. *Journal of Speech, Language and Hearing Research*, *42*, 1482–1498.

Berg, T., & U. Schade. (2000). A local connectionist account of consonant harmony in child language. *Cognitive Science*, *24*, 123–149.

Bernhardt, B. H., & J. P. Stemberger. (1998). *Handbook of phonological development: From the perspective of constraint-based nonlinear phonology*. San Diego, CA: Academic Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. L., & D. I. Slobin. (1982). Rules and schemas in the development and use of the English past-tense. *Language*, *58*, 265–289.

Corina, D. (1994). The induction of prosodic constraints: Implications for phonological theory and mental representations. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules*, pp. 115–145. Amsterdam: John Benjamins.

Daugherty, K., & M. S. Seidenberg. (1992). Rules or connections? The past tense revisited. *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*, pp. 259–264. Hillsdale, NJ: Erlbaum.

de Lacy, P. (2006). *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press.

Dell, G. S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science*, *9*, 3–23.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*, 313–349.

Dell, G. S., C. Juliano, & A. Govindjee. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, *17*, 149–195.

Dell, G. S., & P. G. O'Seaghdha. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review*, *98*, 604–614.

Dressler, W. U. (1979). Experimentally induced phonological paraphasias. *Brain & Language*, *8*, 19–24.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 213–252.

Fodor, J., & Z. Pylyshyn. (1988). Connectionism and cognitive architecture. *Cognition*, *28*, 3–71.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation*, vol. 9, pp. 133–177. New York: Academic Press.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*, 594–621.

Hare, M., & J. Elman. (1992). A connectionist account of English inflectional morphology: Evidence from language change. *In the Proceedings of the 14th Annual Conference of the Cognitive Science Society*, pp. 265–270. Lawrence Erlbaum.

Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes*, *21*, 856–891.

Jaeger, J. J. (2004). *Kids' slips: What young children's slips of the tongue reveal about language development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Jakobson, R. (1941/1968). *Child language, aphasia, and phonological universals*. (A. R. Keiler, trans.) The Hague: Mouton, 1968. (Original: *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell).

James, D. G. H. (2006). *Hippopotamus is so hard to say: Children's acquisition of polysyllabic words*. Unpublished Ph.D. dissertation, University of Sydney, Sydney, Australia.

Jordan, M. (1986). *Serial order: A parallel distributed processing approach*. In Institute for Cognitive Science Report 8604. San Diego: University of California.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations*, part 1, pp. 130–175. Dordrecht, Holland: Foris.

Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*, 589–600.

Laubstein, A. S. (1999). Lemmas and lexemes: The evidence from blends. *Brain & Language*, *68*, 35–43.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.

MacKay, D. G. (1973). Complexity in output systems: Evidence from behavioral hybrids. *American Journal of Psychology*, *86*, 785–806.

Macken, M. A., & D. Barton. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word initial stop consonants. *Journal of Child Language*, *7*, 41–74.

MacNeilage, P. F. (2015). Central tenets of the frame/content theory of evolution and acquisition of speech production. In M. A. Redford (Ed.), *The handbook of speech production*, pp. 353–378. Hoboken, NJ: John Wiley & Sons.

MacWhinney, B., & J. Leinbach. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, *40*, 121–157.

Magnuson, J. S., J. A. Dixon, M. K. Tanenhaus, & R. N. Aslin. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*, 133–156.

Marchman, V. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science*, *21*, 283–304.

Marcus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, & F. Xu. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57* (whole no. 4).

Mason, G. (2015). *Multisyllabic word production of school-age children with and without protracted phonological development within a nonlinear phonological framework*. Unpublished dissertation, University of British Columbia.

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, *12*, 373–418.

McClelland, J. L., & D. E. Rumelhart. (1981). An interactive activation model of context effects in letter perception: Part I: An account of basic findings. *Psychological Review*, *88*, 375–407.

McClelland, J. L., & D. E. Rumelhart. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.

Munson, B., J. Edwards, S. Schellinger, M. E. Beckman, & M. K. Meyer. (2010). Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical Linguistics & Phonetics*, *24*, 245–260.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences*, *23*, 443–512.

Paradis, C., & J.-F. Prunet. (1991). Asymmetry and visibility in consonant articulations. In C. Paradis & J.-F. Prunet (Eds.), *The special status of coronals*, pp. 1–28. San Diego: Academic Press.

Pater, J., & J. A. Barlow. (2003). Constraint conflict in cluster reduction. *Journal of Child Language*, *30*, 487–526.

Petrič, T., & J. P. Stemberger. (2014). Permanent and temporary phonological influences in Slovenian irregular verb production. *Language & Cognitive Processes*, *29*, 470–482.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S., & A. S. Prince. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*, 73–194.

Pinker, S., & M. T. Ullman. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456–463.

Plaut, D. C., & C. T. Kello. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language*, pp. 381–415. Mahwah, NJ: Lawrence Erlbaum Associates, Ltd.

Plunkett, K., & P. Juola. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, *23*, 463–490.

Plunkett, K., & V. Marchman. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, *38*, 43–102.

Plunkett, K., & R. C. Nakisa. (1997). A connectionist model of the Arabic plural system. *Language & Cognitive Processes*, *12*, 807–836.

Pouplier, M. & L. Goldstein. (2005). Asymmetries in the perception of speech production errors. *Journal of Phonetics*, *33*, 47–75.

Rumelhart, D., & J. McClelland. (1986). On learning the past tenses of English verbs. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1, pp. 216–271. Cambridge, MA: Bransford Books.

Scobbie, J., F. Gibbon, W. Hardcastle, & P. Fletcher. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V*, pp. 194–207. Cambridge, UK: Cambridge University Press.

Seidenberg, M. S., & J. L. McClelland. (1989). A distributed, developmental model of visual word recognition and naming. *Psychological Review*, *96*, 523–568.

Shattuck-Hufnagel, S., & D. Klatt. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech errors. *Journal of Verbal Learning and Verbal Behavior*, *18*, 41–55.

Smit, A. B. (1993). Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: Consonant singletons. *Journal of Speech & Hearing Research*, *36*, 533–547.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language*, vol. 1, pp. 143–186. Hove, UK: Lawrence Erlbaum Associates Inc.

Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language*, *28*, 164–188.

Stemberger, J. P. (1990). Wordshape errors in language production. *Cognition*, *35*, 123–157.

Stemberger, J. P. (1991a). Apparent anti-frequency effects in language production: The Addition Bias and phonological underspecification. *Journal of Memory and Language*, *30*, 161–185.

Stemberger, J. P. (1991b). Radical underspecification in language production. *Phonology*, *8*, 73–112.

Stemberger, J. P. (1992a). Vocalic underspecification in English language production. *Language*, *68*, 492–524.

Stemberger, J. P. (1992b). A connectionist view of child phonology: Phonological processing without phonological processes. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications*, pp. 165–189. Timonium, MD: York Press.

Stemberger, J. P. (1993). Vowel dominance in overregularization. *Journal of Child Language*, *20*, 503–521.

Stemberger, J. P. (1994). Rule-less morphology at the phonology–lexicon interface. In R. Corrigan, G. Iverson, & S. Lima (Eds.), *The reality of linguistic rules*, pp. 147–169. Amsterdam: John Benjamins.
Stemberger, J. P. (2004). Neighbourhood effects on error rates in speech production. *Brain & Language*, *90*, 413–422.
Stemberger, J. P. (2007). Children's overtensing errors: Phonological and lexical effects on syntax. *Journal of Memory & Language*, *57*, 49–64.
Stemberger, J. P. (2009). Preventing perseveration in language production. *Language & Cognitive Processes*, *24*, 1431–1470.
Stemberger, J. P. (2013). Perseveration and speech errors in child language production. In P. Brooks, V. Kempe, & J. G. Golson (Eds.), *Encyclopedia of language development*. Thousand Oaks, CA: Sage Publications.
Stemberger, J. P., & B. MacWhinney. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, *14*, 17–26.
Stemberger, J. P., & C. M. Middleton. (2003). Vowel dominance and morphological processing. *Language & Cognitive Processes*, *18*, 369–404.
Stemberger, J. P., & C. Stoel-Gammon. (1991). The underspecification of coronals: Evidence from language acquisition and performance error. In C. Paradis & J.-F. Prunet (Eds.), *The special status of coronals*, pp. 181–199. San Diego: Academic Press.
Stemberger, J. P., & R. Treiman. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, *25*, 163–180.
Stoel-Gammon, C. (1996). On the acquisition of velars in English. In B. Bernhardt, J. Gilbert, & D. Ingram (Eds.), *Proceedings of the UBC international conference on phonological acquisition*, pp. 201–214. Somerville, MA: Cascadilla Press.
Stoel-Gammon, C., & J. P. Stemberger. (1994). Consonant harmony and phonological underspecification in child speech. In M. Yavaş (Ed.), *First and second language phonology*, pp. 63–80. San Diego, CA: Singular Press.
Tyler, A. A., G. R. Figurski, & T. Langdale. (1993). Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. *Journal of Speech and Hearing Research*, *36*, 746–759.
Vihman, M. M. (1978). Consonant harmony: Its scope and function in child language. In J. H. Greenberg (Ed.), *Universals of human language, vol. 2: Phonology*, pp. 281–334. Stanford, CA: Stanford University Press.
Vihman, M. M. (1996). *Phonological development*. Oxford: Blackwell.
Vitevitch, M. S. (2002). The influence of phonological similarity neighbourhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 735–747.
Wagner, K. R. (1985). How much do children say in a day? *Journal of Child Language*, *12*, 475–487.

Part II
Approaches

15
Substance Free phonology

Charles Reiss

15.1 Introduction

Imagine a theory of phonology that makes no reference to wellformedness, repair, contrast, typology, variation, language change, markedness, 'child phonology', faithfulness, constraints, phonotactics, articulatory or acoustic phonetics, or speech perception. What remains in such a phonological theory constitutes the components of the Substance Free phonology (SFP) model I will sketch here. My task thus involves not only justifying the exclusion of all those domains, but also arguing that something remains that is worthy of the name 'phonology'. In support of the latter task, I'll provide some positive examples of recent research in SFP. Of course, the assumption underlying this theory is that there exists a correspondingly narrow object of study in the world, a substance free phonological module of the human language faculty.[1]

To get an idea of what I mean by a substance free theory, let's assume that there are human languages that show word-final obstruent devoicing and languages that do not, but no languages that show word-final obstruent *voicing*. A theory of phonology is *substance free* if it *cannot* capture such apparently true generalizations, and it is not substance free if it can.[2] The fundamental assumption of SFP is that the former type of theory, a 'clean' theory, is preferable to the latter, 'substance abusing', type theory. A phonological theory *should not*, by itself, account for every true generalization about attestable phonological systems. In particular, it should not account for generalizations about statistics of attested or attestable patterns of phonetic substance, even those that are presumed to be absolute, such as the (assumed here) impossibility of final voicing.[3]

The SFP approach accepts the existence of an innate, universal feature set. The features relate, via complex transduction processes, to phonetic substance (sections 2 and 3), but varieties of substance do not have different effects on the computational system – the formal properties of a feature deletion process do not change depending on whether it is +Voiced or −ATR that is being deleted. With this distinction of substantive features and a formal computational system, SFP is not significantly different from what Chomsky (1965, 28) lays out in in *Aspects of a Theory of Syntax*:

> The study of linguistic universals is the study of the properties of any generative grammar for a natural language. Particular assumptions about linguistic universals may

> pertain to either the syntactic, semantic, or phonological component, or to interrelations among the three components.
>
> It is useful to classify linguistic universals as *formal* or *substantive*. A theory of substantive universals claims that items of a particular kind in any language must be drawn from a fixed class of items. For example, Jakobson's theory of distinctive features can be interpreted as making an assertion about substantive universals with respect to the phonological component of a generative grammar. It asserts that each output of this component consists of elements that are characterized in terms of some small number of fixed, universal, phonetic features (perhaps on the order of fifteen or twenty), each of which has a substantive acoustic-articulatory characterization independent of any particular language.
>
> . . . Substantive universals such as these concern the vocabulary for the description of language; formal universals involve rather the character of the rules that appear in grammars and the ways in which they can be interconnected.

In phonology, a formal universal would be the discovery that the phonology of all languages is a complex function, the composition of a strictly ordered set of rules of some well-defined class or some alternative computational system.

Below, I return to the source, nature and number of features, the substantive universals admitted by Chomsky and SFP. For now, note that in *Aspects* Chomsky's characterization of substantive universals has no bearing on the statistical distribution of typological patterns, putative patterns in the course of acquisition or other so-called markedness phenomena. The substantive universals just determine the content of the representations that are the arguments of the computational system. In the SFP view, phonological UG cannot contain a condition that, say, only segments that are voiced and rounded are subject to deletion; or a condition that only, say, round, back, nasal and ATR features can participate in harmony processes – even if this were a true generalization about *languages*, we would not want to encode it in UG as a property of *Language*, the human language faculty.

15.2 Is SPE substance free?

Despite the clarity of *Aspects*, the idea that the phonetic substance, or "intrinsic content", of phonological features should be relevant to the formal component phonological theory was entertained late in *The Sound Pattern of English* (*SPE*) by Chomsky and Halle (1968, 400):

> The problem is that our approach to features, to rules and to evaluation has been overly formal. Suppose, for example, that we were systematically to interchange features or to replace [αF] by [-αF] (where α is +, and F is a feature) throughout our description of English structure. There is nothing in our account of linguistic theory to indicate that the result would be the description of a system that violates certain principles governing human languages. To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made use of the fact that the features have intrinsic content.

Chomsky and Halle are bemoaning the fact that their model developed in the previous 399 pages is "overly formal". The model could easily be used to describe a language with final *voicing*, for example. This call for a theory of markedness in generative phonology is perhaps responsible for inspiring most work in phonology for the last five decades, from the universal processes of Natural Phonology to the universal markedness constraints of Optimality Theory.

Note that Chomsky and Halle seem to be suggesting the pursuit of a theory of markedness that would complement the formal theory of rules they have developed. Thus, even with a theory of markedness, there would remain in *SPE* what we can anachronistically call a formal phonology *module* that would potentially be substance free.

But we don't really have to worry about whether or not the markedness theory alluded to in *SPE* would impinge upon the formal phonology component, since Chomsky and Halle, within the same chapter, point out the futility of pursuing such a model:

> It does not seem likely that an elaboration of the theory along the lines just reviewed will allow us to dispense with phonological processes that change features fairly freely. The second stage of the Velar Softening Rule of English . . . and of the Second Velar Palatalization of Slavic strongly suggests that the phonological component requires wide latitude in the freedom to change features, along the lines of the rules discussed in the body of this book.
>
> *[428]*

In other words, there are rules and rule combinations that effect alternations that are surprising given the "intrinsic content" of phonological features. In yet other words, an adequate model of phonology must be *substance free* – we can't combine the formal theory with a markedness theory that actually constrains it.

15.3 Is Optimality Theory substance free?

In Optimality Theory (OT) (McCarthy and Prince (1993); Prince and Smolensky (1993); Kager (1999); *inter alia*) the evaluation of candidates via the Eval function proceeds without regard to the content of the constraints. Only the ranking of constraints and the relationship of an input form to the candidate forms play a role in applying Eval. This part of the model is substance free.

However, substance is abusively present in other parts of OT, most notably in the content of the constraints: "The basic idea we will explore is that Universal Grammar consists largely of a set of constraints on representational well-formedness, out of which individual grammars are constructed" (Prince and Smolensky (1993, 3)). Since the constraint set Con is assumed to be universal, the model contains specific constraints, like ones that are violated by voiced obstruents in codas. Such constraints will yield final devoicing, when appropriately ranked *vis-à-vis* faithfulness constraints that are violated by mismatches between an input and a candidate output. However, Con does not contain a complementary constraint violated by *voiceless* obstruents in codas. The supposed benefit of this model is that the factorial typology, the set of languages describable by reranking the innate candidate set, matches, in principle, the ones we could find attested. By assumption, we can find languages with final devoicing, and languages without voicing alternations in codas, but no languages with final *voicing*. OT aims to capture such (presumably true) generalizations. This aspect of OT is substance abusing.

Much OT work explicitly appeals to phonetic, physiological and physical factors to explain the inclusion of a given innate constraint in Con. For example, McCarthy and Prince (1995, 88) refer to a constraint *VgV as the "phonologization of Boyle's Law", a law that governs the relationships among volume, pressure and temperature of a gas. Prince and Smolensky (1993) explicitly reject the extreme formalist position of a substance free phonology:

> We urge a reassessment of this essentially formalist position. If phonology is separated from the principles of well-formedness (the 'laws') that drive it, the resulting loss of constraint and theoretical depth will mark a major defeat for the enterprise.
>
> *[216; see also p. 3]*

In other words, OT (at least as represented by most work of its three founders, McCarthy, Prince and Smolensky) advocates building into UG, as constraints in Con, phenomena that have independent explanations via phonetic, physiological and physical factors. As pointed out by John Ohala in various contexts (e.g. 1990) it is not better in science to have two explanations (phonetics *and* phonology) rather than one (*just* phonetics) for a given observation.

Alan Prince (2007, 46) has more recently retreated from seeking justification for OT constraints in phonetic, physiological and physical factors:

> A constraint, in the intended sense, is a principle within a theory and, like any other principle in any other theory, is justified by its contribution to the consequences of that theory. Since OT is a theory of grammar, the consequences are displayed in the grammars predicted and disallowed – 'typological evidence'. A constraint which cannot be justified on those grounds cannot be justified. Further, justifying a constraint functionally (or in any other extrinsic way) can have no effect whatever on its role within the theory. A constraint, viewed locally, can appear wonderfully concordant with some function [ease of articulation, ease of perception, *etc*. – CR], but this cannot supplant the theory's logic or compel the global outcome ('efficiency') that is imagined to follow from the constraint's presence, or even make it more likely.

So, Prince is no longer appealing to grammar-external phonetic, physiological and physical factors to ground constraints. Instead he is proposing that we just posit for Con the constraints that are needed to get the analyses we observe. We posit them because we need them. This *appears* to be good science, basically Occam's Razor. And it appears to be an improvement over Prince's earlier Boyle's Law-type explanations, because there is no appeal to phonetic, physiological and physical factors to explain the ontogeny of constraints. However, this position is even worse than the earlier position.

Instead of merely duplicating in the innate content of Con explanations of typological patterns due to phonetic, physiological and physical factors, Prince's new view contains the same redundancy, but tries to make a virtue of ignoring it. The new approach does not ground constraints, but it leaves the overwhelming phonetic naturalness of many phonological processes and their overlap with OT constraints as an unacknowledged mystery. This is a step backwards that is anticipated by Prince's 'explanation' of markedness presented in Tesar et al. (1999, 305):

> The concept of linguistic markedness, or inherent complexity of structures, has played a significant role in linguistic thought in the twentieth century, though more often as an unformalized perspective or side issue within generative grammar. Optimality Theory rests directly on a theory of linguistic markedness: "marked", or linguistically complex, structures are literally marked as such by the constraints they violate.

This attempt to formalize markedness is reiterated on the next page: "The basic notion of a marked structure is directly built into the theory: a marked structure is one receiving a violation mark by a constraint in Con." Again, if markedness is equated merely with constraint violation 'marks', it remains a mystery why so many of the constraints that assign those marks to structures seem to have good phonetic motivation. Furthermore, as David Odden (p.c.) points out, this definition of markedness destroys the distinction between wellformedness and correspondence/faithfulness constraints that is so central to all versions of OT.

The way forward, in the twenty-*first* century, is to abandon markedness, and to develop a modular theory that distinguishes incidental facts about phonologies (including statistical generalizations about the set of attested ones) from phonological facts (facts about a module of the human mind). The nature of speech perception and sound change drive to a great extent the distribution of patterns we find in the languages of the world, but these distributions are facts about particular phonological systems; they are not facts about phonological UG. Reiterating arguments made elsewhere, I'll try to clarify this distinction in the next section.

15.4 Formedness, a.k.a. markednesslessness

Consider the strings in (1):

(1) a. The cat left.
  b. *Cat the left.

Grammaticality is a relative notion. A string *s* is grammatical with respect to a grammar *G* if and only if *G* generates *s*. The string *s* may be grammatical with respect to *G*, but not grammatical with respect to another grammar *F*. In other words, despite common usage, there is no such thing as an ungrammatical sentence in this technical sense. A string is a sentence with respect to a grammar, or it is not.[4] From this perspective, calling (1a) a grammatical or wellformed sentence is redundant once we know that the grammar in question is mine – it is a sentence of my language; and calling (1b) an ungrammatical or illformed sentence is nonsensical. Sentences are not ill- or well-formed – they are just what grammars generate. Sentences are 'formed' and strings that do not correspond to sentences generated by a grammar are not 'formed'. Of course, it is conceivable that a grammar other than mine generates (1b) but not (1a).

SFP adopts for phonology the same view of the notion of wellformedness presented here for syntax, namely that it is an empty or even misleading notion for understanding grammar. It also rejects the idea that grammaticality is a gradient status. Underlying phonological representations are generated as the output of morphological operations such as concatenation; the resulting underlying forms are fed into a phonological grammar (a complex function); and surface representations result as the output of this function. Those output representations are grammatical in the sense that they are generated by a morphological and phonological grammar. A form like [bunt] might be grammatical as the output of a grammar with (transparent) final devoicing. A form like [bund] will not be ungrammatical or illformed as output of that grammar. It just won't be formed. It won't exist as a surface form of the language.

Calling [bund] 'illformed' represents a basic confusion. The grammar outputs [bunt] as a possible form, so [bunt] is grammatical or *formed*. The form [bund] is a hypothetical form dreamed up by a linguist to demonstrate what is *not* in the intensionally defined set of grammatical forms generated by the grammar in question. There is no reason to ascribe to the grammar in question the property of assigning to [bund] any status whatsoever. We, the linguists, call it 'ungrammatical' to mean 'not generated'. There are an infinite number of things that are not generated by the grammar, for an infinite number of reasons. My simple perspective is that the grammar intensionally generates a set of forms, and any form we linguists pull out of the air that is not in that set is not formed (by that grammar) by implication.

The notion of markedness or wellformedness is fundamental to OT and to many of its predecessors, and that is why these models fall prey to substance abuse. The rhetoric of wellformedness can be quite colorful, with reference to 'conditions' and 'cures' or 'repairs' being commonplace:

(2) Phonological pathology

a. "The main contribution of the OCP is that it allows us to separate out condition and cure. The OCP is a trigger, a pressure for change."

[Yip (1988, 74)]

b. "Repairs have the function of converting phonological configurations marked as illicit by active constraints into licensed ones."

[Calabrese (2005, 75)]

c. "OT takes on a difficulty that held back earlier approaches to naturalness: the *what* is phonetically difficult is not the same as the *how to fix it*."

[Hayes and Steriade (2004, 2)]

Calabrese's reference to "active constraints" allows him to account for crosslinguistic variation in which representations are allowed to surface in a given language and which need to be repaired.

Calabrese discusses the point made by Hale and Reiss (2008, 1998) that it is strange to build constraints into UG against front rounded vowels or ejective stops. For speakers learning languages that have such sounds – say, French and Navaho, respectively – the UG-given gift is misleading, since these sounds do occur. For a speaker of English, such constraints are irrelevant since there is no reason for an English learner to ever posit them – they are not present in the input. We argued that UG, which is supposed to help solve the paradox of language acquisition and explain how kids learn language, should not be full of hints that are at best irrelevant and at worst misleading. Calabrese's response is interesting:

> I agree with [the] claim that the fact that the English child has the "knowledge" that ejective stops are marked is irrelevant from the point of view of the grammar that has been learned. Not so, however, for the Navaho learner where markedness predicts that more effort is needed to learn and produce the complex ejective stops. The point is that for human beings certain actions are more complex than others. Thus, for example, a double backward somersault is more complex than a cartwheel in gymnastics insofar as it requires more complex muscular co-ordinations. Learning how to perform this acrobatic stunt will thus involve a lot of training and effort so that this stunt will be learned only after the easier cartwheel. Once the training is achieved, the backward double somersault is easily performed, albeit still intrinsically complex, by a trained gymnast. Notice that it will be easily lost with the passing of time and that any small health problem will affect its implementation. The same can be said of phonologically marked segments. For the speaker of Navaho, obviously well trained in the pronunciation of this

> language, the ejective stops, although intrinsically difficult, will be easy to pronounce. In contrast, the English speaker, who has never been exposed to ejective stops, will experience problems if exposed to them in not having been trained in their pronunciation. This is what the presence of an active marking statement indicates, and the solutions that speakers will find to the problems posed by segments disallowed by an active marking statement will involve "grammatical" repairs. Resorting to grammatical repairs is the only way speakers have to deal with these segments other than learning how to pronounce them, which means deactivating the relevant marking statement.
>
> *[44]*

Calabrese's discussion can be critiqued from many angles, but most interesting perhaps is the unsubstantiated claim that ejective stops are obviously universally difficult to make (and the implication that Navaho speakers will lose the ability to make them if they don't practice for a long time!); and the further assumption that the presumed physical complexity of the action should correspond to a complexity in the mental representation of the action.

As far as I know, there is no well-developed model of the mental representation of human physical actions. Do we know for example whether such representations can contain recursive structures? We do know that speech in general is quite complex, as illustrated in this nice passage brought to my attention by Veno Volenec:

> Speech is our most human characteristic. It is the most highly skilled muscular activity that human beings ever achieve, requiring the precise and rapid co-ordination of more than eighty different muscles, many of them paired. Even the expertise of the concert pianist pales into relative insignificance beside the intricately co-ordinated muscular vocal skills exercised by a ten-year-old child talking to friends in the school playground. A pianist playing a rapid arpeggio makes about sixteen finger strokes per second, each the product of multiple motor commands to the muscles of the fingers, wrist and arm. Speech is both faster and more complicated (Boomer, 1978). The process of speaking at a normal rate is achieved by means of some 1,400 motor commands per second to the muscles of the speech apparatus (Lenneberg et al., 1967). As children, we take a number of years to acquire the skills of producing and perceiving speech, but once we have learned these abilities, only pathology or accident deprives us of their use.
>
> *[Laver (1994, 1)]*

So, in light of the incredible complexity of all speech, it is not at all obvious that ejectives are more complex than non-ejectives, or that any potential difference would be significant with respect to supposed markedness criteria such as typological prevalence. After all, even the apparently articulatorily simple schwa vowel, made with the mouth in a 'neutral' position, is lacking from many familiar vowel inventories, such as those of Spanish, Italian and Hungarian.

However, for the sake of argument, let's suppose we did have some reason to believe that ejectives are articulatorily more complex than non-ejectives. Would that support Calabrese's position? This appears to be a good example of what Pylyshyn (2003, 2) calls the "the mistake of attributing to a mental representation the properties of what it represents", and the mistake appears to be shared by Hayes and Steriade, whose statement in (2c) suggests that phonetic difficulties have to be fixed by the grammar.

The mental representation of a mouse is not smaller than that of an elephant; the mental representation of a brick is not heavier than that of a feather; the mental representation of the diameter of an atom is not smaller than the mental representation of the diameter of the Milky Way; and even if we demonstrate that producing [p'] is more physically challenging than producing [p], it does not follow that the mental representation of the former is more complex than that of the latter and needs to be fixed. As I become ever more decrepit, the 'passing of time' and 'small health problems' may have degraded my capacity to dunk a basketball or perform a gargouillade, but this is not a matter of mental representation.

It should be obvious that without a notion of wellformedness, there is no sense to the idea of the grammar optimizing output forms in any way. The grammar generates not the best form, but just the form it generates. Other forms are not worse or less optimal; they just are not generated. The idea of wellformedness led OT into what we might call the *fallacy of imperfection* (in response to McCarthy and Prince's (1994) *fallacy of perfection*). For more general critique of the use of constraints in linguistics, see Reiss (2008) and relevant parts of Hale and Reiss (2008).

SFP rejects all repair and optimization approaches to motivating both language-specific and universal phonological computations, because there is no useful sense in which linguistic representations are broken or illformed and need to be fixed. Mental representations either exist or they don't; forms are either generated by a grammar or not. There is no sense in which a mental representation can be illformed or wellformed any more than a molecule can be illformed or wellformed. An existing molecule is compatible with the laws of physics; an 'illformed molecule' that violates the laws of physics is no molecule at all – it does not exist. If a representation exists, it is a possible representation and does not need to undergo repair. Note that even mental representations of physically impossible objects, like the Devil's Triangle or an Escher impossible staircase, are well-formed *qua* mental representations.

SFP rejects language-specific and universal notions of wellformedness for the simple reason that grammars map to outputs from inputs. By invoking a hypothetical form *h* that is not an output of a grammar *G*, we don't magically endow *h* with a status relative to *G*. In sum, the notion of wellformedness, or equivalently, markedness, needs to be rejected to achieve a substance free theory in any linguistic domain. We need to accept markednesslessness, the non-existence of markedness, to progress in phonology.

Mark Hale and I have elsewhere (1998, 2008) written about the irrelevance of so-called child phonology to discussions about markedness and complexity – let's just point out here that the work on infant speech perception (e.g. Werker, 1995) guarantees that Navaho children will distinguish ejectives from non-ejectives, and thus they have the capacity to represent the distinction. Calabrese's suggestion that "a lot of training and effort" is required for kids to have *representations* of ejectives is basically a reversion to Piagetian views of cognitive development as dependent on sensorimotor experience (see Karmiloff-Smith (1992) for critical yet sympathetic discussion of Piaget). Once again, it should be pointed out that the SFP perspective is not new, since Chomsky made the same point in 1964, despite his close contact with Jakobson whose work (e.g. Jakobson, 1971) shows that he clearly believed that observation of child speech was revealing of deep generalizations about phonology.

Chomsky (1964, 39), commenting on a conference paper on child 'phonology', notes that there is:

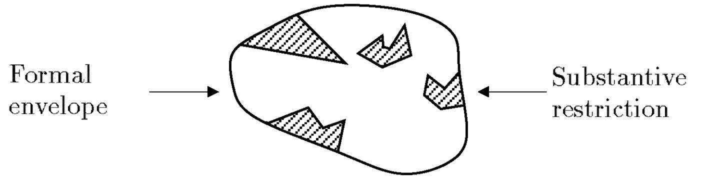
> a general tendency to oversimplify drastically the facts of linguistic structure and to assume that the determination of competence can be derived from description of a corpus by some sort of sufficiently developed data-processing techniques. My feeling is that this is hopeless and that only experimentation of a fairly indirect and ingenious sort can provide evidence that is at all critical for formulating a true account of the child's

grammar (as in the case of investigation of any other real system). . . . I make these remarks only to indicate a difficulty which I think is looming rather large and to which some serious attention will have to be given fairly soon.

It is gratifying to see that Neil Smith, one of the most prominent and careful scholars of child 'phonology', has most recently re-evaluated the nature of the evidence to conclude, consistent with the SFP approach, that child speech does not bear on phonological UG in the way that most work in the field has assumed, such as OT work on the emergence of the unmarked in child speech. Smith now favors the view that "the major determinant of children's divergent productions is performance rather than the competence" (Smith, 2010, 103).

Ron Kaplan (1987, 346–7) raises Chomsky and Halle's concession for the need for a formal model unconstrained by substance to the status of a methodological principle for linguists, illustrated with his jocular, but useful, Figure 2, reproduced here in (3), along with his discussion:

(3) The shape of a linguistic theory (Kaplan's Figure 2)



A formal theory may have a relatively smooth outline . . . [t]hen you start taking chunks out of it . . . because you claim that no human language or grammar has such and such a property. . . . It's a mistake to carry premature and unjustified substantive hypotheses into our computational and mathematical work, especially if it leads to mathematically complex, even if more restrictive, theories. . . . [W]e should be wary of the seduction of substance.

Chomsky and Halle were uncharacteristically weak in their conclusions about markedness, to such an extent that their own students could later treat issues of substance as core, universal components of phonology in the form of OT markedness constraints against, for example, front round vowels and ejectives. However, there were explicit calls for substance free phonology even around the time of publication of *SPE*, such as Fudge (1967, 26): "phonologists (above all, generative phonologists) ought to burn their phonetic boats and turn to a genuinely abstract framework"; and Hellberg (1978, 157):

A certain phonological rule may be perfectly well statable in terms of distinctive features. And more than that: it may to some extent have a good phonetic plausibility [but] it need not have any direct phonetic motivation whatsoever synchronically today.

Fudge (2006) even recognizes an explicit substance free approach in the work of the glossematicians Hjelmslev and Uldall, whose methodology was a general scientific approach in which "substance must be excluded from consideration" and one that aims to "set up hypotheses about what must be the abstract formal system capable of accounting for the data" (Fudge 2006: 1440). Considerations of substance, such as a feature system, are added on top of the fundamental work of characterizing a formal system.

So, why can't we just treat innate OT constraints on a par with the innate features of *SPE* (and most OT), as a complement to the formal system of ordered rules or Eval or some alternative? To reiterate, if phonetic grounding (like Boyle's Law) is supposed to play a role in determining the content of the constraints, then we have a duplication of explanation in phonetics and the content of Con; if phonetic grounding is irrelevant (Prince's more recent view) then the duplication is an unacknowledged mystery.

I have suggested (Reiss, 2008) that what Prince and Smolensky call the "principles of well-formedness (the 'laws')" of phonology are actually just heuristics that we develop through our experience as linguists looking at lots of languages. In other words, they are not part of the ontology of phonology. For example, looking at a new language, we typically assume that it is likely that a sequence like [akra] has a syllable boundary before the stop-liquid cluster, rather than between the two consonants. This is because we seem to believe, rightly or wrongly that most languages 'maximize onsets' in such cases and leave the first syllable without a coda.

However, both syllabifications are found, for example, in the Ancient Greek dialects (Steriade, 1982). It may be useful to assume that the more common syllabification is present in a new, unfamiliar language until there is evidence to the contrary; and the guess will turn out to be correct more often than not, if our professional intuitions have any basis. However, we must take care not to confuse our intuitions concerning what happens often with the actual nature of the system under study. Based on our experiences and expectations, we apply our intuitions in attempting to solve the problems involved with analyzing data, but there is no reason to expect that these intuitions directly reflect the nature of the actual mental grammar constructed by a learner. The intuition that heavy things fall faster than light things is very useful when someone drops something from a window, but the intuition needs to be transcended to understand the workings of gravity. Heuristics are used by the analyst to make useful guesses about data, and guesses can be wrong. This is why OT constraints need to be violable – they reflect the fallibility of our guesses. If this perspective is valid, then the great innovation of OT, the violability of its constraints, represents a basic category error – the constraints correspond to linguists' intuitions about what processes are common, not to the ontology of phonological UG. In this case, all the formal work on OT will have been for nought. Violable constraints will go the way of the ether.

15.5 Contrast

Another notion to dismiss as we construct SFP is *contrast*. This is sure to be yet another unpopular move: "Contrast . . . is one of the most central concepts in linguistics" (Dresher 2009). It is important, before we proceed, to stress the distinction between (a) the components of our proposed model of the Human Phonological Faculty – our object of study; and (b) our methods for making hypotheses about that object – our sources of evidence, the heuristics we use. To use fancy names, we need to distinguish *phonological ontology* (which asks *What are the components of universal and language specific phonologies?*) from *phonological epistemology* (which asks *How do we make discoveries and justify claims about universal and language specific phonologies?*).

Let's look at this distinction with respect to *contrast*. Contrast is *not* part of the ontology of SFP. However, practitioners of SFP have no qualms about referring to minimal pairs or the fact that French has a contrast between oral and nasal vowels. This is because we use our knowledge of the existence of semantic contrast and minimal pairs as a basis for hypotheses about the phonological content of elements in a French-type lexicon. We use minimal pairs epistemologically to justify our beliefs about French phonological ontology.

For SFP the notion of contrast is used as a heuristic, a tool to help us discover and understand phonology, whereas other work in phonology accepts contrast as part of the ontology of phonology. Work that appeals to contrast as part of the ontology of phonology falls along a spectrum of sophistication; however, most work fails to even consider the distinction I am making, or else explicitly rejects it. Some work, such as that of the Toronto school (e.g. Dresher (2009) and related work), is of course quite complex in its argumentation for building contrast into the model of grammar. Other work, such as that of Flemming (2004, 1995), is refreshing in the honesty of its unabashed functionalism. Flemming (2004, 232) is interested in "investigating the general character of the constraints imposed on phonology by the *need to minimise confusion* [which] is hypothesised to derive from the communicative function of language" [emphasis added]. This perspective sets Flemming's work, including his Maintain Contrast OT constraints, apart from almost all modern phonological research in the generative tradition, which generally eschews overtly functionalist reasoning, and sometimes rejects it explicitly:

> Since language is not, in its essence, a means for transmitting [cognitive] information – though no one denies that we constantly use language for this very purpose – then it is hardly surprising to find in languages much ambiguity and redundancy, as well as other properties that are obviously undesirable in a good communication code.
>
> *[Halle (1975, 528)]*

Lexical and structural ambiguity, as well as the existence of neutralization rules (even if the neutralization is incomplete!), all illustrate Halle's point. In SFP, we carefully distinguish the nature of language from the use to which we sometimes put it.

It is not possible to evaluate here every appeal to the role of contrast in grammar; however, it is simple to appreciate in light of Halle's statements the connection between (ontological) contrast and another idea that is anathema to SFP, *functionalism*. The intended sense of functionalism here is the idea that insight into grammar can be gleaned from consideration of the fact that language is used for communication. It is apparent that the notion of contrast, as used in phonology, relates to the capacity of a phonological difference to communicate a difference in meaning. Many phonological discussions of contrast since Jakobson have borne the taint of functionalism and a failure to distinguish epistemology from ontology.

In this brief overview, I cannot survey the relatively sophisticated attempts to build contrast into phonology, for example, attempts to show that phonological processes can be sensitive to the distinction between contrastive and non-contrastive features in a language. However, given its relationship to functionalism, and given the failure to clearly distinguish epistemological from ontological questions, I suggest that the null hypothesis should be that contrast is *not* relevant to phonology. Papers such as Odden (2017) and Reiss (2017) discuss the relationship between contrast and markedness in greater detail.

15.6 Banishing phonotactics

A staple of phonological discussion since Chomsky and Halle (1965) is the differential evaluation of forms like [blɪk] and [bnɪk] by English speakers. Neither form corresponds to a word (or morpheme) in a speaker's lexicon, yet the former is judged to be a possible word and the latter not a possible word. The robustness of such judgments is taken to reflect speakers' knowledge of the phonotactics of their language, generalizations about what sound combinations are existent/wellformed. Since phonotactic knowledge is knowledge about sound patterns, it is often assumed to be part of phonology. Also, some aspects of phonotactics are clearly due to the effects of phonological rules – a language with a rule that always assimilates /n/ to a following /p/ or /k/ will not have sequences like [np] or [nk] in output forms (barring opaque rule interaction).

OT and other approaches to phonology assume that phonotactic patterns must be accounted for by the phonological grammar. The OT solution is interestingly elegant in that the surface inventory of segments and the relations among segments, the phonotactics, emerge from the same constraint ranking that accounts for alternations, without the need for morpheme structure constraints on the form of lexical entries or other statements.

I suggest that the arguments for assuming that phonotactics is part of grammar are weak, and that the null hypothesis should be that phonotactic judgments reflect metalinguistic awareness, similar to awareness of rhyme or metrics for poetic composition and parsing. Hearkening back to the early arguments for rationalist generative grammar (e.g. Katz and Bever, 1976), I suggest that phonotactic judgments are like acceptability judgments in reflecting many factors, and thus not necessarily a good indication of grammaticality status.

A first argument against putting too much weight on phonotactic judgments is that there is no reason to think that speakers' judgments are always valid. North American speakers, for example, take a lot of convincing to judge the flaps of *rider* and *writer* as identical. They also will assert that no words begin with a [pt] cluster, even when they clearly pronounce *potato* with such a cluster. It is actually not even clear to me what the 'correct' answer is – if there *is* a vowel in the output representation of *potato* that is so reduced in speech as to be unidentifiable in a spectrogram, is there or is there not a phonotactic ban on initial [pt] clusters? Speaker's phonotactic judgments are colored by orthography, morphophonemic alternations, experience with other languages and accents, and so on.

Despite their potential lack of reliability, we can sometimes find surprising accuracy in judgments of phonotactic patterns in the definite *absence* of grammatical knowledge. Suppose a group of monolingual, literate, English-speaking non-linguists are presented orally with forms like [pumehana] and [bɛzvzglɛndnɨ] and they are asked which one sounds like Polish and which like Hawaiian. I imagine the subjects would pretty much all agree in their judgments. Informally, we could say that their judgments are 'correct'. We will have thus demonstrated some kind of phonotactic 'knowledge' in the absence of a grammar, since our (hypothetical) subjects speak neither Polish nor Hawaiian. If we asked for judgments about Polish words *vs.* Russian words, or Hawaiian *vs.* Samoan words, the subjects might perform worse, and we expect performance to be tied to past exposure to the relevant languages. In this context, we would expect the same (English-speaking) subjects, who obviously have lots and lots of exposure to English data much like their own dialects, to have very strong judgments about English phonotactics. And this is what we find. Therefore, there is no reason to assume that strong, and often correct, phonotactic judgments reflect grammatical knowledge. People have such judgments about languages for which they clearly have no grammar. So, SFP, as a theory of grammar, has no problem with failing or refusing to account for phonotactic 'knowledge'.

A final reason to reject phonotactics as part of phonology is that Frisch et al. (2000) found that phonotactic judgments are gradient, suggesting that such judgments are non-grammatical in nature. Interestingly, the authors of the study drew a very different conclusion, claiming that the results on phonotactic judgments showed that grammar itself, and grammatical wellformedness (not just speaker *judgments* of grammaticality), is gradient. This strikes me as throwing out the categorical baby of discrete symbolic computation *instead of* throwing out the phonotactic bathwater that I am happy to see go.

SFP ignores phonotactic judgments – they are not phenomena that the *grammar* must describe. Even when accurately reflecting facts about occurring sound sequences, such judgments may be drawing on a wide array of factors, including frequency of sequences in the lexicon, token frequency in spoken language surface forms, and generalizations that reflect the history of the language. For example, if a sound change in the history of *L* assimilated /n/ to /m/ before /p/, then the absence of /np/ sequences in *L* may be a static fact about the lexicon of *L*, about which speakers may have judgments. In the absence of alternations, there is no reason to expect the computational system to account for such a static generalization.

Work on phonotactics like Daland et al. (2011, 198) asserts that phonotactic principles like the Sonority Sequencing Principle are known to be "synchronically active in speakers' grammars", based on what they call "sonority projection effects" concerning how speakers judge various non-occurring consonant clusters. I won't go into details here, but the important quality of these projection effects is that "the offending clusters are systematically and *equally* absent from speakers input, and yet speakers appear to *differentiate* some clusters as less well-formed than others" (Daland et al. 2011, 198). I suggest that the phonotactic literature is making the same mistakes that Katz and Bever (1976) pointed out in the generative semantics literature of the seventies:

> generative semantics has distorted grammar by including within its goals those of a complete theory of acceptability.
>
> . . . The issue between the rationalist and the empiricist conception of the domain of grammar is an empirical one. Our estimate of the evidence at present is that it heavily confirms the rationalist strict separation of grammatical phenomena in the traditional sense from extragrammatical phenomena. . . . [W]e have shown that the rationalist program can not only deal with the phenomena brought up but does so in a more satisfactory way. Moreover, as we have already indicated, the generative semanticists' criterion leads to a theory that rapidly becomes a study and compilation of everything. But a compilation of everything is a science of nothing: the advantage of the rationalist program, then, is that by distinguishing different contributions to linguistic behavior, explanation in terms of appropriate principles becomes possible in each case.
>
> *[p. 58 ff.]*

SFP, in excluding phonotactics from consideration, in *rejecting* the idea that "any phenomenon systematically related to cooccurrence is *ipso facto* something to be explained in the grammar" (Katz and Bever, 1976, 58), is in the rationalist tradition of generative grammar. The old arguments for this approach seem to hold.

15.7 Features and rules

As pointed out above, the main tenet of SFP is that phonetic substance is not relevant to the formal properties of phonological computation – issues like how rules compose and the formal properties of rules. Strictly speaking, then, a substance free model has no automatic implications for the status of substantive universals. In practice, however, we find a strange situation. On the one hand, SFP is strongly nativist, with full acceptance of a universal, innate feature set. On the other hand, there are other phonologists who adopt the substance free *label*, but who interpret it as a rejection of universal features with consistent phonetic correlates and acceptance of arbitrary rules. To further muddy the waters, there is work that purports to argue against SFP, but in fact argues against views about features that have nothing to do with SFP. Rather than sort out all these matters, I attempt here to lay out the basis for the views favored in my own model that uses the term ‘substance free’ phonology, namely SFP. A good overview of various uses of the phrase ‘substance free phonology’ can be found in Blaho (2008), many of which adopt a view of features at odds with that of SFP.

In brief, I argue in this section the following three points:

1. phonology is epistemologically prior to phonetics
2. features can’t be posited on the basis of rules
3. features must be innate

This whole section recaps a long tradition in generative grammar of defending the rationalist perspective over an empiricist perspective that resurfaces repeatedly. In addition to influencing much of the early work on rationalism from a generative perspective, Chomsky (1965) and Chomsky (1966) trace similar ideas much farther back. As far as I can tell, the logical arguments presented by Chomsky and the others have never been seriously challenged.

15.7.1 The priority of phonology

Most of the discussion on innateness and rationalism *vs.* empiricism is concerned with syntax and semantics – it is really hard to think that features like Plural or Accusative are actually in the signal. Phonologists, however, typically can’t stop themselves from lapsing into thinking that they are working with sounds. In this context, let’s take a lesson from a *rationalist phonetician*, a hero of the SFP school, Robert Hammarberg. Hammarberg (1976) leads us to see that for a strict empiricist, the somewhat rounded-lipped *k* of *coop* and the somewhat spread-lipped *k* of *keep* are very different.[5] Given their distinctness, Hammarberg makes the point, obvious yet profound, that we linguists have no reason to compare these two segments unless we have a paradigm that provides us with the category *k*. Our phonological theory is logically prior to our phonetic description of these two segments as ‘kinds of *k*’. So, our science is rationalist. As Hammarberg also points out, the same reasoning applies to the learner – only because of a pre-existing built-in system of categories used to parse can the learner treat the two ‘sounds’ as variants of a category: “phonology is logically and epistemologically prior to phonetics” (Hammarberg 1976: 354). Phonology provides equivalence classes for phonetic discussion.

Hammarberg’s discussion is grounded in the philosophy of early generative grammar and general philosophy of science:

> Chomskian linguistics is explicitly anti-empiricist, and all indications are that current philosophy of science is moving toward a rejection of the empiricist programme ((Fodor, 1968, pp. xiv *ff)*). A key feature of the new programme is exactly a reevaluation of the concept of observation. Observations are now held to be judgments, and these judgments are made in terms of the criteria provided by the paradigm. Thus the taxonomy of a discipline is to be regarded as imposed from above, rather than emerging from below, i.e., rather than emerging in the form of brute facts before the unprejudiced eyes or ears of the researcher. The relevance of this to the study of phonetics and phonology should be obvious: the concept of the segment, which is indispensable to phonetics and phonology, is a creature of the paradigm, not of the raw data.
>
> *[Hammarberg (1976, 354)]*

Hammarberg (1981, 266) revisits the "Kantian claim that objects conform to our modes of cognition", again drawing on Chomsky, as well as on modern physics:

(4) "The 'furniture of the world' does not come prepackaged in the form of individuals with properties, apart from human intervention: [e]ither the analysis provided by the cognitive system that we might call 'common sense understanding' or the more self-conscious idealizations of the scientist seeking to comprehend some aspect of physical or mental reality."

[Chomsky (1980, 218–9); cf. also Chomsky (1975)]

(5) "The doctrine that the world is made up of objects whose existence is independent of human consciousness turns out to be in conflict with facts established by experiment."

[d'Espagnat (1979, 158)]

The point of all this is that phonological categories can't be learned from phonetics, since there can't be any phonetics without a pre-existing phonology. It is the intervention of our cognitive system, our 'cognoscitive powers' (a term Chomsky (2000a) adopts from seventeenth-century philosophy), that packages sounds, say, into syllables, segments, sentences and so on. These categories belong to the science of phonology as well as the object of the science of phonology (the human phonological faculty). As Hammarberg (1976) says, "It should be perfectly obvious by now that segments do not exist outside the human mind." But they are not fictions: "there would be little value in such an approach. Science aims for a theory of the real, and to base one's descriptions and generalizations on a fictional taxonomy could only lead to one's theories being fictional as well" (354).

15.7.2 Features are innate

SFP adopts fully Hammarberg's realist, rationalist views, and this ties in with general discussions of innateness, beyond phonology. I won't revisit here arguments for the *discrete, binary* phonological categories understood in terms of discrete, binary feature values – let's assume that phonological representations consist of data structures built from valued features, like, say, +Nasal and the like. Hammarberg's arguments reflect a vast literature on the topic, all leading to the conclusion expressed by Jackendoff (1990, 40): "In any computational theory, 'learning' can consist only of creating novel combinations of primitives already innately available. This is one of the fundamental arguments of Fodor (1975), and one that I accept unconditionally."

As Fodor (1975, 82) puts it:

> Trivially, one cannot use the predicates that one is learning in order to learn the predicates that one is using. . . . It follows immediately that not all the languages one knows are languages one has learned, and that at least one of the languages which one knows without learning is as powerful as any language that one can ever learn.

More straightforwardly, whatever the features, the primitives of phonological representation, are, they have to be innate.

Chapter 1 of Chomsky (1965) contains a detailed discussion of the rationalist tradition supporting this nativist perspective, and many other serious scholars, including Fodor and

Jackendoff cited above, have presented it as a logical necessity. Chomsky (1980, 45–46) explains that without such innate abilities/constraints/categories, we would not be able to do much of anything:

> Were it not for this endowment, individuals would grow into mental amoeboids, unlike one another, each merely reflecting the limited and impoverished environment in which he or she develops, lacking entirely the finely articulated and refined cognitive organs that make possible the rich and creative mental life that is characteristic of all individuals not seriously impaired by individual or social pathology – though once again we must bear in mind that the very same intrinsic factors that permit these achievements also impose severe limits on the states that can be attained; to put it differently, that there is an inseparable connection between the scope and limits of human knowledge.

In other words, without innate features, we would not be able to parse input at all, and of course, we would have no way of explaining how people exposed to different input can sometimes arrive at identical knowledge states.

In this context, it is distressing to see that the arguments and conclusions of the rationalist bases of generative grammar have been somewhat cavalierly ignored in the phonological literature in the latest revival of hyper-empiricism. A wide array of scholars have asserted that there is no need for innate features, because features can be discovered or constructed on the basis of induction over the input to the learner. For example, Archangeli and Pulleyblank (2015, 2) tell us to "See Mielke, 2004 [PhD thesis published as Mielke (2008) – CR] on why features cannot be innately defined, but must be learned." However, Mielke's thesis and book do not mention the arguments for the logical necessity of innate representational primitives given by Fodor, Jackendoff, Hammarberg or anyone else. None of these are mentioned in the book at all. Mielke (2008, 27) asserts that "Chomsky and Halle's assumption that distinctive features are innate is treated in subsequent literature as if it were a conclusion", but Mielke is ignoring the centuries of discussion on the topic that is more general than phonology – the acceptance of a universal innate feature set is a specific conclusion based on a general argument made by linguists, philosophers and psychologists. Where Mielke does look beyond phonology, he restricts himself to syntax, and concludes: "Most of the evidence for UG is not related to phonology, and phonology has more of a guilt-by association status with respect to innateness" (34). This is hardly a sufficient refutation.

Hall (2014) characterizes the SFP view well: Hale and Reiss "assume that features are innate and universal, and have substantive phonetic content", but goes on to state that he and Elan Dresher (2015) have both offered rebuttals of the innateness view. Dresher's claim is unambiguous: "There is a growing consensus that phonological features are not innate, but rather emerge in the course of acquisition" (165). Dresher's position is not completely antinativist. He proposes that "a fixed innate list of phonological features has been problematic on empirical grounds, and is not conceptually necessary because there is an innate mental mechanism for creating distinctive features in the course of language acquisition" (2016). This mechanism is said to rely on "an auditory system that allows us to make certain sound discriminations", suggesting that features can be derived from the basic categories of auditory perception.

Such a theory, one that can derive specific features from a general concept of 'feature' and the powers of auditory perception, is in principle a better theory than one that posits the richer innate structure of models like *SPE* and SFP. However, the model seems to fail in light of the problem of the lack of invariance (Appelbaum, 1996), as well as the vast acoustic differences among speakers when saying 'the same thing'. It also appears to be at odds with the results in infant speech perception and the need for parsing input into representations for learning to even begin, as appreciated by Fodor and others. It seems unavoidable to posit higher level equivalence classes for speech perception to even take off in acquisition. In other words, we need innate features.

The SFP position on the necessity of innate features is not something we discovered or adopted lightly. It follows logically from arguments extending at least to Kant, as the above discussion of Hammarberg suggested. To my knowledge, the arguments have never been refuted.

These issues are not restricted to speech perception, of course. Pylyshyn (1984, 13), in a discussion of the stimulus independence of cognitive behavior generally, summarizes the situation in visual perception: "virtually no candidate physical properties (for example, particular physical features) are either necessary or sufficient for a person perceiving some situation in a certain way." He surveys the various ways in which perception of a Necker cube can be induced. These include stimuli that vary colors and shapes of points in a static display, presentation of an image behind a moving slit or a moving image behind a fixed slit, and even the effect of combining two images in a random dot stereogram, each half of which has no independent properties that correspond to the lines of the cube. Drawing on a linguistic notion, Pylyshyn (1984, 13) states that perceptual "regularities that exist are to be found among perceived (cognitively described) properties or what Pike (1967) calls *emic* properties, not among objective (physically described) or *etic* properties".

15.7.3 Can rules tell us what the features are?

Some authors who reject innateness use the term 'substance free' to refer to the idea that features are not substantive universals, but are rather induced from patterns in the learner's input. Hall (2014, 2–3) gives an excellent characterization of this view:

> In this view [also called 'substance free' – CR], features are not universal or innate, but rather are induced by the learner. Featural representations are assigned on the basis of phonological behaviour, not acoustic or articulatory substance (although phonological properties often do happen to correlate with phonetic ones). This allows for rules that are maximally formally elegant, even when they are phonetically unnatural. As Blaho (2008: 2223) puts it, "Features are indicators of the way members of an inventory behave, but they don't necessarily have any consistent phonetic characteristics even within the same system." Likewise, in Emergent Feature Theory, features do not necessarily have any content beyond identifying "the segments that do X" (Mielke 2008: 99).

The problem with this perspective is that there is no way to induce patterns from phonological behavior without an innate feature system with which to parse the input. In Fodor's terms, you need an innate language to learn another language. Isac and Reiss (2013) provide an accessible demonstration of this logic using toy grammars based on playing cards (adapted from Hale and Reiss (2008) and Hale and Reiss (2003b)).

Another problem with these approaches is that they assume that it is obvious what a rule is. Suppose the input provides evidence for alternations that delete word-final /a/ and word-final /t/. Are we to expect the learner to posit a feature that makes {a,t} a natural class of segments? Of course not.

The SFP approach, laid out in Bale et al. (2014) and Bale and Reiss (forthcoming) is basically that of *SPE*: the target and environment segments of a rule are sets of segments, each member of which is a superset of some set of features – the segments form a natural class. If we intersect the features of /a/ and /t/, we get a natural class that will contain lots of other segments that do *not* delete word-finally. In SFP, this result tells us that there must be two separate deletion rules, one for /t/ and one for /a/. You can't induce the features from the rules without a theory of rules, and a theory of rules relies on natural classes. You can't figure out which segments "do X" without the means to identify segments and identify X.

Let's now turn to a familiar example where the observable evidence suggests that the triggers of a rule are *not* a natural class, but where a deep analysis provides a solution in which the rule is in fact based on natural classes. The regular English plural marker is underlying /-z/. The single segment of this morpheme surfaces as [s] when occurring after a root ending in one of these segments: {p,t,k,θ,f}.[6] Now, the participation of the segments {p,t,k,θ,f} as triggers suggests to the observer that these segments should form a natural class. However, they do not according to all feature systems in use – any class that contains these segments should contain [s] and [ʃ], too.

A simple solution to this problem is to say that the rule inserting a vowel between two coronal stridents (as in *bushes*) comes before and bleeds the devoicing rule. In other words, the devoicing rule *intensionally* can be formulated to apply to sounds specified −Voice, but the extensionally defined set of observable segments that trigger the rule is {p,t,k,θ,f}, which is not a natural class in English. It is impossible to decide if rules are natural or unnatural with respect to an innate feature set if we depend only on "observable distributional regularities". There is a vast literature demonstrating that the "relation between a phonemic system and the phonetic record . . . is remote and complex" (Chomsky, 1964, 38). Neither segments nor rules are observable. They are instead the outcome of an analysis (by linguist or learner) in which the "essential properties underlie the surface form" (Katz and Bever, 1976, 12).

The extent to which this basic logic of rules and natural classes is misunderstood is quite surprising. Even authors interested in the formal properties of grammars get the relationship backwards. For example, Kornai (2008, 29), assuming a language with a segment inventory that includes at least [p,t,b,d], says that "Phonologists would be truly astonished to find a language where some rule or regularity affects *p*, *t*, and *d* but no other segment" – presumably he has *b* in mind. The SFP view is that there cannot be such a rule because the natural class defined by features would force us to posit *two* rules, not one. Without innate substantive features, we can't determine what the rules are, and we can't determine what the segments are. Features are the primitive units that allow the learner to parse the signal into segments and determine if various alternations can be collapsed into a single rule.

In the absence of any engagement with the logical arguments of Fodor and Chomsky and others, the assertion that features can be induced from the rules cannot be taken seriously. SFP continues to assume, therefore, an innate set of phonological primitives.

This is not to say that the exact set of features proposed in the literature, or anything close to the number of features proposed, is close to accurate (see Hale et al. (2007) for arguments that standard proposals are way too low). But we must accept that there is an innate feature set.

I wish to reiterate that the SFP phonology perspective is not new. The following quote illustrates the extent to which the idea that generalizations of grammar are observable on the surface was rejected in syntax, at least, in the early days of generative grammar:

> From the general intellectual viewpoint, the most significant aspect of the transformationalist revolution is that it is a decisive defeat of empiricism in an influential social science. The natural position for an empiricist to adopt on the question of the nature of grammars

> is the structuralist theory of taxonomic grammar, since on this theory every property essential to a language is characterizable on the basis of observable features of the surface form of its sentences. Hence, everything that must be acquired in gaining mastery of a language is "out in the open"; moreover, it can be learned on the basis of procedures for segmenting and classifying speech that presuppose only inductive generalizations from observable distributional regularities. On the structuralist theory of taxonomic grammar, the environmental input to language acquisition is rich enough, relative to the presumed richness of the grammatical structure of the language, for this acquisition process to take place without the help of innate . . . principles about the universal structure of language. Rationalists, on the other hand, find the taxonomic theory uncongenial because, for them, the essential properties of language underlie the surface form of sentences and are thus unobservable in the sense in which atoms are unobservable.
>
> *[Katz and Bever (1976, 12)]*

The real English devoicing rule, one which is featurally natural (it has to be or it would not be a rule for SFP) is not observable. Instead, we observe a phonetically unnatural set of triggers. The solution comes from a sophisticated analysis in terms of ordered rules that cannot be read from the signal.

Before leaving the issue of the innateness of features, it is worthwhile pointing out that the innateness model appears to be the only one compatible with the well-established experimental results of infant speech perception found by Janet Werker and her collaborators (see Werker (1995) for an overview). It beggars belief that the sensitivity to every possible phonemic distinction which Werker finds in infants is unrelated to an innate capacity for phonological representation.

15.7.4 Rejecting a particular model of features

SFP accepts the existence of innate substantive universals, but it does not have to accept all versions of innate representational schema for features. One version that is particularly odious to the SFP perspective is Feature Geometric models that mimic (somewhat) the structure of the vocal tract with the effect of sneaking substance back into the computational system. McCarthy (1988, 84), in an exposition of Feature Geometry, states: "The goal of phonology is the construction of a theory in which cross-linguistically common and well-established processes emerge from very simple combinations of the descriptive parameters of the model." For example, "Assimilation is a common process because it is accomplished by an elementary operation of the theory – addition of an association line" (86). With hindsight, it is apparent that this argument is invalid. The human phonological faculty is only one factor determining the set of attested phonological systems. There is no obvious way in which its properties could determine the common-ness of attested patterns. Such reasoning, extended to syntax, would lead to a theory that makes *do*-support simple to model, because English-type grammars are quite widespread. A more promising view of the goal of phonology and linguistic theory in general is "to abstract from the welter of descriptive complexity certain general principles governing computation that would allow the rules of a particular language to be given in very simple forms" (Chomsky, 2000b, 122). In Bale and Reiss (forthcoming); Reiss and Shen (ms.); and Bale et al. (2016), we explore the limits of basic set theoretic notions for accounting for the behavior of phonological entities. The argument does not rely on what is common versus rare, but rather follows standard scientific method of expanding the model's power only when necessary.

There are many, many issues that remain unsolved with respect to the nature of featural representation and transduction between features and, say, sound. However, these cannot be addressed here. Some of them, like the problem of the lack of invariance, are profound, longstanding issues in the study of speech perception (see Appelbaum (1996) for discussion) that should probably be considered as separate from the computational system of phonology proper. This problem is part of the issue of the abstract nature of even the substantive universals accepted by SFP. In SFP the old question of whether features are primarily acoustic or primarily articulatory does not make sense – they are mental representations, primitives of mental data structures, in the sense of Gallistel and King (2009), each with a complex transduction relation to input and output systems. Given that even non-linguistic auditory perception is subject to illusions and thus sometimes non-veridical (Bregman, 1990; Reiss, 2007), it can't be the case that phonological features would be straightforwardly related to acoustics or articulation.

15.7.5 How rich is phonological UG?

SFP adopts the view of Gallistel and King (2009) that mental representations in general must consist of a hierarchy of data structures that are ultimately composed of some set of atomic symbols. The effect of having such a taxonomy of symbols is that a relatively small number of lower level symbols can be combined into a very large number of higher level symbols. This idea is consonant with Chomsky's (2007, 4) point that "the less attributed to genetic information (in our case, the topic of UG) for determining the development of an organism, the more feasible the study of its evolution" – it is simpler to study the evolution of a simple system than a complex one. Gallistel and King's discussion of how to use combinatorics to deal with combinatoric explosion is illustrated for phonology in Reiss (2012) where I walk through the simple math that shows that a UG with just four binary features and the option of having underspecified segments allows for 2.4 septillion languages.

So, the combinatorics makes the point that we don't need a lot of features in UG to get a lot of descriptive power from UG, but obviously we need more than four phonological features. It is common to find in the phonological literature that twenty-five or so features is a reasonable number (recall that Chomsky suggested "perhaps on the order of fifteen or twenty" in *Aspects*). The SFP perspective is that this is an arbitrary number with no current justification – we are nowhere close to knowing the exact number of features, and it is silly to think that any of the features discussed in the literature are, without a doubt, correctly identified and individuated. Hale et al. (2007) argue that it is preferable to increase somewhat the number of features in our models of UG, rather than posit language specific phonetic implementation rules. The combinatoric explosion that differentiates models of UG with twenty-five or fifty or 100 features is unfathomable, but irrelevant. If we sensibly think of this all in terms of an intensionally defined UG, the differences among these numbers of features grows merely linearly, and they are all of the same order of magnitude.

15.8 Some SFP

So, what's left?

Much discussion in the literature of the formal properties of phonological computation is either compatible with SFP, or, if it proves superior to current SFP proposals, should be incorporated into SFP in their stead. Every phonological theory contains a substance free component, and many of the ideas from this work have informed SFP. Because of its rejection of so much recent work, SFP can appear reactionary, but the intention is maintenance of the good arguments concerning the rationalist basis of generative linguistics with a critical evaluation of longstanding notions like markedness and contrast in phonology, all accompanied by technical and analytic contributions.

The impressive body of work by Jeff Heinz and his colleagues and students on characterizing phonology from the perspective of formal language theory is probably the most influential recent work in formal phonology that is substance free in the SFP sense. I refer the reader directly to this work, including Heinz and Idsardi (2013); Heinz (2010); Chandlee et al. (2014); Chandlee and Heinz (2012); and Chandlee and Koirala (2014). While this work is refreshing in its explicitness and in the coherence of this budding research community, its ultimate worth for a linguist in the generative tradition will depend on the significance of formal language theoretic results to generative concerns and the extent to which the work can generate new insights into phonological phenomena, and not just formalize potentially problematic aspects of traditional phonological work, such as phonotactics.

I exclude from consideration as relevant to SFP the even larger literature on the formal properties of OT because, despite the sophistication of these discussions (e.g. work by Jason Eisner, Jason Riggle, Giorgio Magri, Alan Prince, Bruce Tesar and many others), in all instantiations OT ends up being substance abusing by virtue of the content of Con and the notion of wellformedness embedded in the very idea of optimality – there can be no optimal form unless the other forms are less optimal, less wellformed, as discussed above. Recall that “Optimality Theory rests directly on a theory of linguistic markedness” (Tesar et al., 1999, 305). If markednesslessness is correct, this foundation is rotten. It remains to be seen if all that impressive mathematical work on OT formalism can be salvaged.[7]

Much work in syllabification and metrical phonology either is untainted by substance abuse, or else can be fairly easily detoxified and recast in substance free terms. An excellent body of work compatible with SFP developed from Raimy’s (2000a) research on precedence relations and linearization in phonology. It is worth pointing out that Raimy (2000b) was able to demonstrate, *contra* the claims of McCarthy and Prince (1995), that a derivational model can in fact handle supposed cases of under- and over-application which were paraded as demonstrating the failure of derivational models and the need for parallel constraint satisfaction models like OT. One SFP approach to reduplication, Reiss and Simpson (2017), inspired somewhat by Raimy’s work, does without the notions of markedness, Base and Copy, and Correspondence, notions which play an important role in much functionalist and surface-oriented OT work in morphology and phonology (see Hale et al. (1998)).[8]

The Search and Copy models developed somewhat in parallel (and with useful cross-fertilization) by Andrew Nevins (2010), on the one hand, and Shen (2016); Mailhot and Reiss (2007); and Samuels (2011), on the other, are quite similar. Nevins’ approach, in its reliance on contrast and markedness, is not fully compatible with SFP. However, it would probably be useful to extract the best formal aspects from the two traditions and combine them into a better substance free model of locality in phonological rules.

In the following paragraphs I will describe a few more aspects of my own work that exemplify my vision of SFP. Obviously, my discussions of these topics are all subject to criticism on empirical and theoretical grounds, but I hope that they are at least consistently substance free.

Quantifiers in phonology In Reiss (2003) I argued that, to handle so-called anti-gemination and anti-anti-gemination phenomena, phonological rules must be able to compute identity and non-identity between segments. This work is an outgrowth of Odden (1988), a phonologist whose interest in formal issues has generated a lot of SFP-compatible work. (See Bakovic (2006) for an interesting, basically substance free OT critique of my claims.) I argued that such computation is best expressed *via* the power of first order quantificational logic. These conditions, which I expressed with the existential and universal quantifiers, are part of language specific rules. In these quantificational computations over sets of features the particular nature of the features play no role, so the rules are substance free. It should be obvious that many other models, including OT, may also make use of similar quantificational logic, for example to determine whether a constraint against geminates is violated.

Phonological acquisition In Hale and Reiss (2003a, 2008, 1998) we attempt to model phonological acquisition without recourse to markedness, with an appreciation of the contribution of performance issues to the nature of children's output, and in a manner consistent with findings from infant speech perception studies. This work accepts the existence of innate substantive features, but the modeling of the acquisition process proceeds in a substance free fashion, based on simple set theoretic operations – there is no appeal to markedness and no attempt to account for children's superficial speech output, in light of all the evidence from covert contrasts (Gibbon, 1990; Gibbon and Scobbie, 1997) and comprehension studies that demonstrate the sophistication and detail of their representations. We obviously deny the validity of 'the emergence of the unmarked' (TETU) phenomena widely discussed in the OT literature on phonological acquisition and computation, such as Struijke (2014).

Operations in phonology In Bale et al. (2014), we deconstruct the arrow of traditional phonological rules and argue that the arrow corresponds to at least two different operations: set subtraction and unification. We make use of this distinction to revive and formalize an old idea from Poser (1993, 2004): 'feature changing rules' should be understood as deletion (which we model via set subtraction) followed by insertion (which we model via unification).

Types and underspecification in phonology Building on the work in Bale et al. (2014), Bale et al. (2016) explore the ramifications of treating segments as sets of valued features. For example, we demonstrate that it is possible to extensionally target only the fully underspecified segment that contains no features, {}, by writing a rule that intensionally targets *all* segments.

This small sample merely suggests the range of issues that are relevant to the SFP perspective. A good example of a substance free topic that has engendered healthy controversy and interesting observations about what various formal systems are capable of is the question of whether there can be *polarity rules*, rules that turn αF into $-\alpha$F, for some feature. Work in OT and various rule-based frameworks has addressed this quintessentially formal problem (see Moreton (1999); Bale et al. (2014); Fitzpatrick et al. (2004) for a sample).

15.9 Conclusions

Three core questions for SFP are the following:

- What kind of data structures are phonological representations?
- What is a possible phonological rule?
- What kind of complex function is a phonology?

Future research will explore these questions, as well as guide exploration of issues in phonological acquisition and the interface of phonology with other modules of grammar.

In this sketch, I have tried to clarify several issues that surround the use of the term ‘substance’ in phonology, especially in the phrase ‘substance free phonology’. The focus has been on laying out my own idiosyncratic model and laying claim to the phrase as a proper name, Substance Free Phonology. Importantly, the freedom from substance in SFP refers only to the nature of the computational system. SFP accepts, indeed embraces, the innate substantive entities of, say, *SPE*, that is the idea of innate universal features. The exact number of features needed for UG and the exact phonetic correlates of the features remain questions for future research.

Although the issues are poorly understood, we entertain the possibility that the *specific* substance of the features is universal across spoken languages because of the universality of the interface of the substantive primitives with the human transduction systems. If the same innate feature set interfaced with a different transduction system, the phonetic correlates would be different. Perhaps this is what happens in signed languages – the same innate feature set interfaces with the visual and manual motor systems. This is very speculative, but has no bearing on the substance free nature of phonological computation, as understood in SFP.

Despite rejecting a tremendous array of topics that are typically considered to lie within the purview of phonological theory — topics like contrast, typology, child speech and markedness — I suggested that there are rich opportunities to understand the nature of phonological computation and representation from a purely formal perspective. There is a lot of work to be done on the nature of representations as data structures; on the kinds of operations that apply to these structures, such as unification and set deletion, as well as quantificational operations and Search and Copy procedures.

Finally, I hope to have demonstrated that SFP takes seriously the arguments for rationalism and other philosophical foundations of generative grammar, including the competence–performance distinction, that have been ignored or rejected (sometimes with uncanny parallels to mistakes of the past) without argument in much recent work. Whatever its failings, I hope the SFP approach rests on the right foundations.

15.10 Further reading

My book-length work with Mark Hale contains many of the ideas presented here. However, it has become apparent that in many instances we were not sufficiently clear. The current chapter attempts to improve and extend some of the ideas in that book:

Hale, Mark, and Charles Reiss. 2008. *The phonological enterprise*. Oxford University Press.

My forthcoming textbook with Alan Bale attempts to present Substance Free Phonology to students:

Bale, Alan, and Charles Reiss. 2018. *Phonology: A formal introduction*. MIT Press.

The following works, ranging from general cognitive science down to general linguistics and phonology, specifically, have helped my understanding in many areas related to this chapter.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
Gallistel, Charles R., and Adam Philip King. 2011. *Memory and the computational brain: Why cognitive science will transform neuroscience*, volume 6. Hoboken, NJ: Wiley-Blackwell.
Hammarberg, Robert. 1976. The metaphysics of coarticulation. *Journal of Phonetics* 4: 353–363.
Hellberg, Staffan. 1978. Unnatural phonology. *Journal of Linguistics* 14.2: 157–177.

Katz, Jerrold J., and Thomas G. Bever. 1976. The fall and rise of empiricism. In *An integrated theory of linguistic ability*, ed. Thomas G. Bever, Jerrold J. Katz, and D. Terence Langendoen, 11–64. Thomas Y Crowell Company.

Pylyshyn, Zenon W. 1986. *Computation and cognition: Toward a foundation for cognitive science*. The MIT Press.

Notes

1 Most of the ideas in this chapter are either discussed in my joint work with Mark Hale, including Hale and Reiss (2008), or grew out of this joint work or more recent work and discussion with Alan Bale, Dana Isac and others.

2 This example is chosen for expository purposes and relies on an assumption of binary features. With privative features, it will sometimes be possible to model such typological asymmetries because of the inventory of the representational primitives, not because the computational system treats the primitives differentially.

3 This discussion parallels early generative work on the contrast between grammaticality and acceptability, for example in Katz and Bever (1976). See the discussion in Chapter 1 of Hale and Reiss (2008) for the purview of UG, and the analogy with grammaticality in Chapter 11 of Isac and Reiss (2013).

4 Equivalently, a string is a sentence in a language or it is not. Of course, we really are interested in *structures* associated with these strings, since sentences are not just strings.

5 Of course, any two tokens of, say, the rounded-lipped *k* are also quite distinct from each other by any number of physical, observable metrics, but let's ignore that and assume we recognize two entities, 'rounded *k*' and 'spread *k*'.

6 Let's say that this is an observable fact, even though it assumes a lot of filtering of the data – 'plurality' and the property of being 'regular' are not in the signal, for example.

7 There remain other apparent problems with OT. The original arguments against intermediate representation appear empty (Karttunen, 1998), and although current versions of OT have reintroduced derivations, with intermediate representations (see work in Harmonic Serialism and Stratal OT, for example), the claim that OT is a two-level model, and the idea that this is somehow desirable, persist.

8 Note that the Output–Output Correspondence and Uniform Exponence Constraints of much OT work parallel the transderivational constraints of the generative semantics literature, strengthening the comparisons between the approaches made above:

> Thus, there must be rules that apply not to individual derivations, but to classes of derivations. In short, transderivational constraints are required, since there are cases where the well-formedness of one derivation depends on certain properties of other, related derivations.
>
> *(Lakoff, 1973)*

References

Appelbaum, Irene. 1996. The lack of invariance problem and the goal of speech perception. In *The 4th International Conference on Spoken Language Processing*, Philadelphia, PA, USA, October 3–6, 1996. ISCA. URL www.isca-speech.org/archive/icslp_1996/i96_1541.html.

Archangeli, Diana, and Douglas Pulleyblank. 2015. Phonology without universal grammar. *Frontiers in Psychology* 6:1–12.

Bakovic, Eric. 2006. Antigemination, assimilation and the determination of identity. *Phonology* 22:279–315.

Bale, Alan, Maxime Papillon, and Charles Reiss. 2014. Targeting underspecified segments: A formal analysis of feature changing and feature filling rules. *Lingua* 148:240–253.

Bale, Alan, and Charles Reiss. 2018. *Phonology: A formal introduction*. MIT Press.

Bale, Alan, Charles Reiss, and David Ta-Chun Shen. 2016. Sets, rules and natural classes: {} *vs*. [ ]. Submitted.

Blaho, Sylvia. 2008. The syntax of phonology: A radically substance-free approach. Doctoral Dissertation, Universitetet i Tromsø.

Boomer, Donald S. 1978. The phonemic clause: Speech unit in human communication. In *Nonverbal behavior and communication*, ed. Aron W. Siegman and Stanley Feldstein, 245–262. Hove: Psychology Press.

Bregman, Albert S. 1990. *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Walter de Gruyter.

Chandlee, Jane, and Jeffrey Heinz. 2012. Bounded copying is subsequential: Implications for metathesis and reduplication. In *Proceedings of the Twelfth Meeting of the Special Interest Group on Computational Morphology and Phonology*, 42–51. Association for Computational Linguistics.

Chandlee, Jane, Adam Jardine, and Jeffrey Heinz. 2014. Learning repairs for marked structures. In *Poster presented at the Annual Meeting of Phonology*. MIT.

Chandlee, Jane, and Cesar Koirala. 2014. Learning local phonological rules. In *Proceedings of the 37th Penn Linguistics Conference*.

Chomsky, Noam. 1964. Formal discussion in response to W. Miller and S. Ervin. In *The Acquisition of Language*, ed. Ursula Bellugi and Roger Brown, 35–39. Chicago: University of Chicago Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1966. *Cartesian linguistics*. New York: Harper & Row.

Chomsky, Noam. 1975. *Reflections on language*. New York: Pantheon Books.

Chomsky, Noam. 1980. *Rules and representations*. New York: Columbia University Press.

Chomsky, Noam. 2000a. Internalist explorations. In *New horizons in the study of language and mind*, 164–194. Cambridge, MA, US: Cambridge University Press.

Chomsky, Noam. 2000b. Language as a natural object. In *New horizons in the study of language and mind*, 106–133. Cambridge, MA, US: Cambridge University Press.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. Uli Sauerland and Hans-Martin Gärtner, volume 89, 1–24. Berlin: Mouton de Gruyter.

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 & Row.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis, and Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28:197–234.

d'Espagnat, Bernard. 1979. The quantum theory and reality. *Scientific American* 241:158–181.

Dresher, Bezalel Elan. 2009. *The contrastive hierarchy in phonology*, volume 121. Cambridge, UK: Cambridge University Press.

Dresher, Bezalel Elan. 2015. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41:165–181.

Dresher, Bezalel Elan. 2016. Reply to comment. URL http://facultyoflanguage.blogspot.com/2016/05/the-return-of-behaviorism.html.

Fitzpatrick, Justin, Andrew Nevins, and Bert Vaux. 2004. Exchange rules and feature-value variables. In *3rd North American Phonology Conference, Concordia University*, Montréal, Québec.

Flemming, Edward. 1995. Phonetic detail in phonology: Toward a unified account of assimilation and coarticulation. *Coyote Papers, Proceedings of the Arizona Phonology Conference, Features in Optimality Theory* 5:1–12.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In *Phonetically based phonology*, ed. Bruce Hayes, Robert Martin Kirchner, and Donca Steriade, 232–276. Cambridge: Cambridge University Press.

Fodor, Jerry A. 1968. *Psychological explanation: An introduction to the philosophy of psychology*. A Random House Study in Problems of Philosophy. New York: Random House.

Fodor, Jerry A. 1975. *The language of thought*. Cambridge, MA: Harvard University Press.

Frisch, Stefan A., Nathan R. Large, and David B. Pisoni. 2000. Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42:481–496.

Fudge, E. C. 1967. The nature of phonological primes. *Journal of Linguistics* 3:1–36.

Fudge, E. C. 2006. Glossematics. In *Encyclopedia of language and linguistics*, second edition, ed. Keith Brown. Oxford: Elsevier.

Gallistel, C. R., and Adam Philip King. 2009. *Memory and the computational brain: Why cognitive science will transform neuroscience*. Chichester, West Sussex, UK: Wiley-Blackwell.

Gibbon, Fiona. 1990. Lingual activity in two speech-disordered children's attempts to produce velar and alveolar stop consonants: Evidence from electropalatographic (epg) data. *British Journal of Disorders of Communication* 25:329–340.

Gibbon, Fiona, and James M. Scobbie. 1997. Covert contrasts in children with phonological disorder. *Australian Communication Quarterly*, 13–16.

Hale, Mark, Madelyn Kissock, and Charles Reiss. 1998. What is output? Output-output correspondence in OT phonology. In *Proceedings of the sixteenth west coast conference on formal linguistics*, ed. Emily Curtis, James Lyle, and Gabriel Webster, 223–236. Stanford, CA: CSLI.

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. 2003a. The subset principle in phonology: Why the tabula can't be rasa. *Journal of Linguistics* 39:219–244.

Hale, Mark, and Charles Reiss. 2003b. 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, USA.

Hall, Daniel Currie. 2014. On substance in phonology. In *Proceedings of the 2014 annual conference of the Canadian Linguistic Association*.

Halle, Morris. 1975. Confessio grammatici. *Language* 51:525–535.

Hammarberg, Robert. 1976. The metaphysics of coarticulation. *Journal of Phonetics* 4:353–363.

Hammarberg, Robert. 1981. The cooked and the raw. *Journal of Information Science* 3:261–267.

Hayes, Bruce, and Donca Steriade. 2004. Introduction: The phonetic bases of phonological markedness. In *Phonetically based phonology*, ed. Bruce Hayes, Robert Martin Kirchner, and Donca Steriade, 1–33. Cambridge: Cambridge University Press. URL www.loc.gov/catdir/description/cam032/2003055722.html.

Heinz, Jeffrey. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41:623–661.

Heinz, Jeffrey, and William Idsardi. 2013. What complexity differences reveal about domains in language*. *Topics in Cognitive Science* 5:111–131.

Hellberg, Staffan. 1978. Unnatural phonology. *Journal of Linguistics* 14:157–177.

Isac, Daniela, and Charles Reiss. 2013. *I-language: An introduction to linguistics as cognitive science*, 2nd edition. Oxford: Oxford University Press, USA.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Jakobson, Roman. 1971. Kindersprache, Aphasie und allgemeine Lautgesetze. In *Selected writings*. The Hague: Mouton.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

Kaplan, Ron M. 1987. Three seductions of computational psycholinguistics. In *Linguistic theory and computer applications*, ed. P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson, and P. Bennett, 149–188. London: Academic Press.

Karmiloff-Smith, Annette. 1992. *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Karttunen, Lauri. 1998. The proper treatment of optimality in computational phonology: Plenary talk. In *Proceedings of the International Workshop on Finite State Methods in Natural Language Processing, FSMNLP'09*, ed. Kemal Oflazer and Lauri Kattunen, 1–12. Stroudsburg, PA, USA: Association for Computational Linguistics.

Katz, Jerrold J., and Thomas G. Bever. 1976. The fall and rise of empiricism. In *An integrated theory of linguistic ability*, ed. Thomas G. Bever, Jerrold J. Katz, and D. Terence Langendoen, 11–64. New York: Thomas Y Crowell Company.

Kornai, András. 2008. *Mathematical linguistics*. London: Springer.

Lakoff, George. 1973. Some thoughts on transderivational constraints. In *Issues in linguistics: Papers in honor of Henry and Renee Kahane*, ed. Braj B. Kachru et al., 442–452. Urbana: University of Illinois Press.

Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.

Lenneberg, Eric H., Noam Chomsky, and Otto Marx. 1967. *Biological foundations of language*, volume 68. New York: Wiley.

Mailhot, Frederic, and Charles Reiss. 2007. Computing long-distance dependencies in vowel harmony. *Biolinguistics* 1.1:28–48.

McCarthy, John J., and Alan S. Prince. 1993. Prosodic morphology 1: Constraint interaction and satisfaction. *Technical Report# 3, Rutgers University Center for Cognitive Science*.

McCarthy, John J., and Alan S. Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts occasional papers in linguistics*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbancyk, volume 18, 249–384. Amherst, MA: GLSA, University of Massachusetts at Amherst.

McCarthy, John J., and Alan S. Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of the North East Linguistic Society*, ed. Mercè Gonzàlez, volume 24, 333–379. Amherst, MA: GLSA Publications. Available on Rutgers Optimality Archive, ROA-13.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Moreton, Elliott. 1999. Non-computable functions in Optimality Theory. Ms., University of Massachusetts, Amherst. Rutgers Optimality Archive ROA-364.

Nevins, Andrew. 2010. *Locality in vowel harmony, volume 55 of linguistic inquiry monographs*. Cambridge, MA: MIT Press.

Odden, David. 1988. Anti antigemination and the OCP. *Linguistic Inquiry* 19:451–475.

Odden, David. 2017. Markedness in substance-free and substance-dependent phonology. In *Arguments of substance: Beyond markedness*, ed. Bridget D. Samuels, 1–22. Amsterdam and Philadelphia: John Benjamins.

Ohala, John. 1990. The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology 1: Between the grammar and physics of speech*, ed. John Kingston and Mary Beckman, 258–275. Cambridge: Cambridge University Press.

Pike, Kenneth L. 1967. *Language in relation to a unified theory of the structure of human behavior*, 2nd rev. Berlin: Mouton & Co.

Poser, W. J. 1993. Are strict cycle effects derivable? In *Studies in lexical phonology*, ed. Sharon Hargus and Ellen M. Kaisse, volume 4, 315–321. San Diego: Academic Press.

Poser, W. J. 2004. *On the status of Chumash sibilant harmony*. Ms., University of Pennsylvania.

Prince, Alan. 2007. The pursuit of theory. In *The Cambridge handbook of phonology*, ed. P. V. De Lacy, 33–60. Cambridge: Cambridge University Press.

Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar (technical report)*. New Brunswick, NJ: Rutgers Center for Cognitive Science.

Pylyshyn, Zenon W. 1984. *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Pylyshyn, Zenon W. 2003. *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.

Raimy, Eric. 2000a. *The phonology and morphology of reduplication*, volume 52. Berlin: M. de Gruyter.

Raimy, Eric. 2000b. Remarks on backcopying. *Linguistic Inquiry* 31:541–552.

Reiss, Charles. 2003. Quantification in structural descriptions: Attested and unattested patterns. *The Linguistic Review* 20:305–338.
Reiss, Charles. 2007. Modularity in the 'sound' domain: Implications for the purview of universal grammar. In *The Oxford handbook of linguistic interfaces*, ed. Gillian Ramchand and Charles Reiss, 53–80. Oxford: Oxford University Press, USA.
Reiss, Charles. 2008. Constraining the learning path without constraints, or the OCP and NoBanana. In *Rules, constraints and phonological phenomena*, ed. Bert Vaux and Andrew Nevins, 252–301. Oxford: Oxford University Press.
Reiss, Charles. 2012. Towards a bottom-up approach to phonological typology. In *Towards a biolinguistic understanding of grammar: Essays on interfaces*, ed. A. M. Di Sciullo, 169–191. Amsterdam: John Benjamins Publishing Company.
Reiss, Charles. 2017. Contrast is irrelevant in phonology: A simple account of Russian /v/ as /V/. In *Arguments of substance: Beyond markedness*, ed. Bridget D. Samuels, 23–46. Amsterdam and Philadelphia: John Benjamins.
Reiss, Charles, and David Ta-Chun Shen. ms. Contour spread: Encoding precedence in tonal representation and beyond. Ms. Concordia U.
Reiss, Charles, and Marc Simpson. 2017. Reduplication as projection. *Revue roumaine de linguistique Presented at GLOW*.
Samuels, Bridget D. 2011. *Phonological architecture: A biolinguistic perspective*. Oxford Studies in Biolinguistics. Oxford: Oxford University Press.
Samuels, Bridget D., ed. 2017. *Arguments of substance: Beyond markedness*. Amsterdam and Philadelphia: John Benjamins.
Shen, David Ta-Chun. 2016. Precedence and search: Primitive concepts in morpho-phonology. Doctoral Dissertation, National Taiwan Normal University.
Smith, Neil. 2010. *Acquiring phonology: A cross-generational case-study*, volume 124. Cambridge: Cambridge University Press.
Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.
Struijke, Caro. 2014. *Existential faithfullness: A study of reduplicative TETU, feature movement and dissimulation*. New York: Routledge.
Tesar, Bruce, Jane Grimshaw, and Alan Prince. 1999. Linguistic and cognitive explanation in Optimality Theory. In *What Is Cognitive Science*, ed. Ernest Lepore and Zenon Pylyshyn, 295–326. Oxford: Blackwell.
Werker, Janet. 1995. Exploring developmental changes in cross-language speech perception. In *An invitation to cognitive science*, ed. Daniel N. Osherson and Lila R. Gleitman, 2nd edition, volume 1, 87–106. Cambridge, MA: MIT Press.
Yip, Moira. 1988. The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19:65–100.

16
The phonology of sign languages

Jordan Fenlon, Kearsy Cormier, and Diane Brentari

16.1 Introduction

Compared to spoken language phonology, the field of sign language phonology is a young one, having begun in the 1960s together with research into sign languages generally. Before this point, linguists often dismissed the academic study of sign languages as manual representations of spoken languages (e.g., Bloomfield, 1933) or as iconic wholes lacking any internal structure. However, since Stokoe's (1960) seminal work, sign language linguists have demonstrated that, as with spoken languages, sign languages have sublexical structure that is systematically organised and constrained. In addition, sign languages also stand in stark contrast to spoken languages because they are produced in the visual-gestural modality and therefore the articulators involved in phonological organisation are different. Within this chapter, we provide an introduction to the field of sign language phonology and a selective overview of contributions to date. We also highlight key areas that have attracted much debate amongst sign language linguists such as the development of phonological models, the effect of modality on phonology, and the relationship between sign language and gesture. In section 16.4, we describe new contributions to the field which have the potential to further illuminate our understanding of sign language phonology in the future. Our description will be centred around two unrelated sign languages: American Sign Language (ASL) and British Sign Language (BSL), though many of the patterns here have been described for other sign languages as well. This chapter's concluding note emphasises that in order to understand phonology, one must consider sign languages.

16.2 Definitions

In this section, we briefly outline sign language phonology and key terms that have been used to refer to the organisation of signs at this level of the language.

16.2.1 The sign language lexicon

Sign language interaction is made up of different types of signs. These different types have been exemplified in models of the sign language lexicon proposed by many researchers

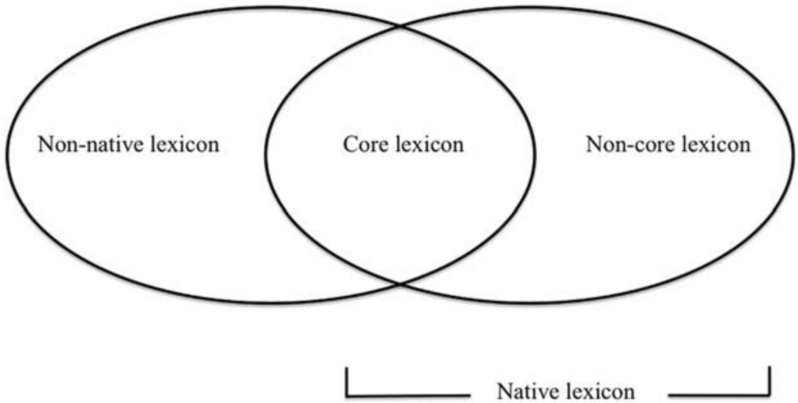


Figure 16.1 Sign language lexicon
(adapted from Brentari and Padden, 2001)

(e.g., Johnston and Schembri, 1999; Brentari and Padden, 2001; Cormier et al., 2012); see Figure 16.1 for a model adapted from Brentari and Padden (2001).

Here the lexicon is divided into the following components: the core lexicon, the non-core lexicon, and the non-native lexicon. Signs in the core lexicon are described as being composed of meaningless sublexical units with a highly conventionalised form and meaning association; these are the signs you would typically expect to see listed in a sign language dictionary. Much of the work on phonological theory concerning sign languages has been based on this component of the lexicon. Signs from the non-core lexicon are, in contrast, made up of meaningful units and typically refer to so-called classifier constructions or signs involving sequences of constructed action (Cormier et al., 2012). Finally, the non-native lexicon refers to fingerspelled sequences. Fingerspelled sequences represent a form of borrowing where different configurations of the hand are each associated with a letter from the corresponding spoken language's alphabet in order to spell out a word. Importantly, signs from the non-core and the non-native lexicon differ in their structural properties from signs in the core lexicon in terms of handshape inventories as well as in the application of phonological constraints and rules (Eccarius, 2008). In the following section, we describe how signs are organised at the phonological level focusing on the core lexicon. We refer to signs from other components of the lexicon in section 16.3.3 when we consider the relationship between signs from the non-core lexicon and gesture, a prominent and current area of enquiry.

16.2.2 The core lexicon: the phonological architecture of the sign

It is widely acknowledged in the sign language literature that the parameters of handshape, place of articulation 'POA' (or location), movement, and orientation play a significant role at the phonological level in a similar way to the spoken language properties of place of articulation, manner, and voicing. In the BSL sign DANGER, the parameters specified are: the flat hand for *handshape*, the ipsilateral side of the forehead for *place of articulation*, and a short repeated movement contacting the forehead for *movement*. *Orientation*, which is interpreted here as the relationship between the active hand and the POA, is the radial side of the hand (i.e., the part of the hand that contacts the POA). Justification for the feature units within each parameter stem from their ability to show contrasts. For example, the BSL sign GAY differs from BSL UNSURE along the handshape dimension alone (BSL GAY has only an extended thumb with all fingers closed, whilst BSL UNSURE has an extended thumb with all fingers extended). In Figure 16.2, pairs of contrasts along each parameter in BSL are provided.

Within the core lexicon, the parameters of handshape, location, movement, and orientation are sometimes viewed as meaningless sublexical elements (e.g., there appears to be no iconic motivation for the handshape in the signs PAPER or DANGER; these are arbitrary sublexical elements that are contrastive in BSL). Several phonological models have been proposed to account for a sign's underlying representation and the organisation of these parameters. Within this section, an overview of the general organisation of the sign

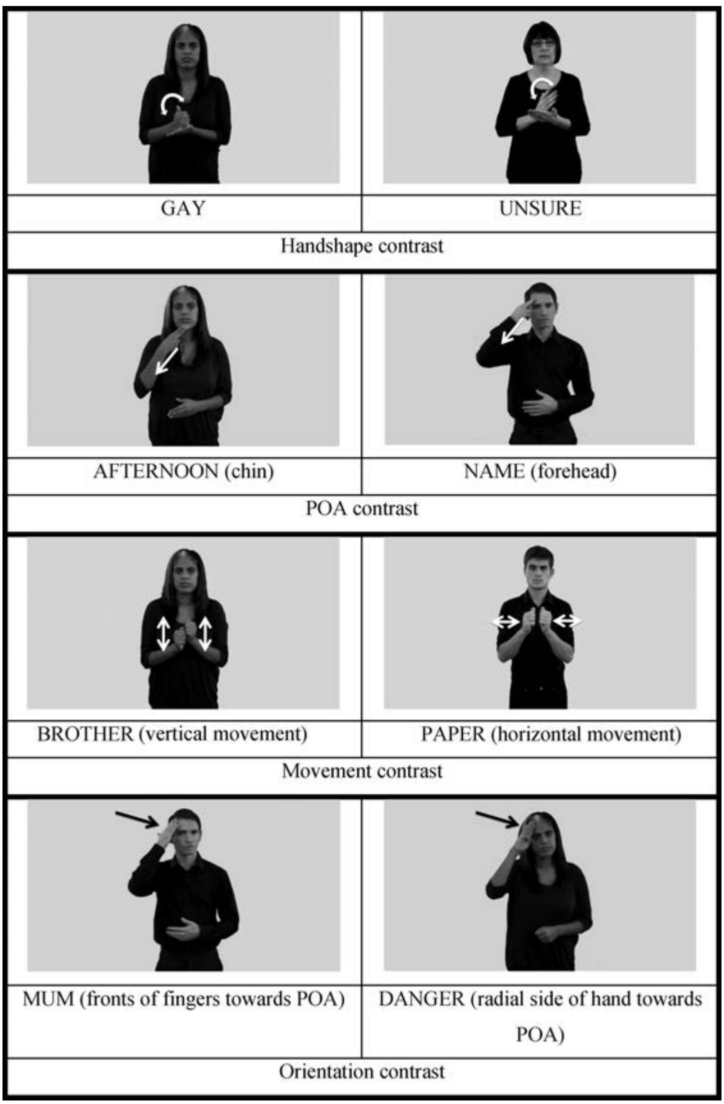


Figure 16.2 Sign pairs with handshape, location movement, and orientation contrasts in BSL (stills from BSL SignBank (Fenlon et al., 2014))

according to the Prosodic Model is provided (see Figure 16.3). We refer again to other phonological models in section 3.1.

The Prosodic Model follows Dependency Theory (Anderson and Ewen, 1987; van der Hulst, 1993; van der Hulst and van de Weijer, *this volume*) in that each node is maximally binary branching, and each branching structure has a head (which is more elaborate) and a dependent (which is less elaborate). In the following sub-sections, we describe how the parameters of handshape, POA, orientation, and movement are represented within the Prosodic Model. These sub-sections will refer briefly to the class nodes of the feature hierarchy.

16.2.3 The core lexicon: Inherent Features

A closer look at the Inherent Features structure within the Prosodic Model is provided in Figure 16.4. The Inherent Features structure branches into the parameters of handshape and POA (location); each will be discussed in turn.

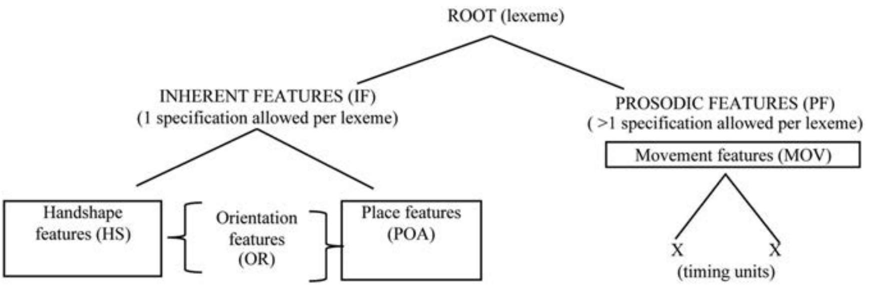


Figure 16.3 Overview of the Prosodic Model

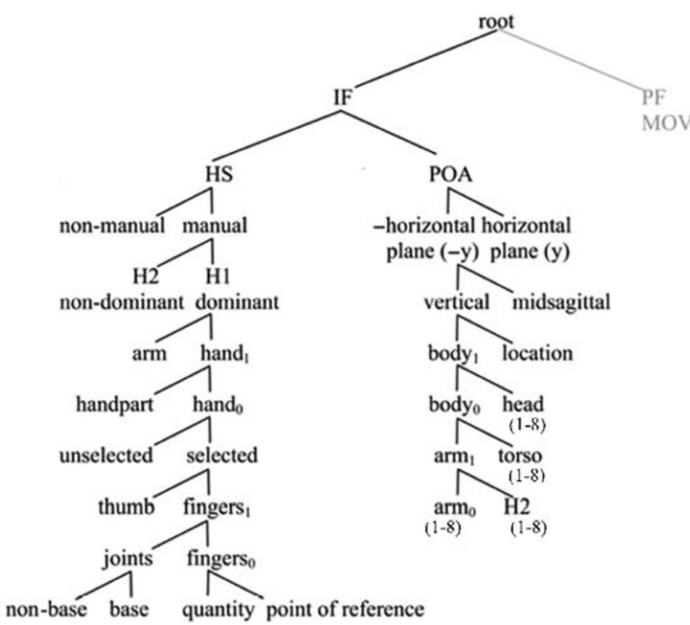


Figure 16.4 Inherent Features structure

16.2.3.1 Handshape

The handshape parameter is specified within the Inherent Features structure. Beginning at the top-most node, the active articulator is first specified, which is typically the arms and hands. In some cases, a sign may only use non-manual articulators (e.g., the head, the face, and/or the body) but these types of signs are relatively rare in signed languages. The manual node then branches into the dominant (H1) and non-dominant (H2) hands. If the sign is two-handed, it will have both H1 and H2 features. If the sign is one-handed, it will only have H1 features. These features include which fingers are 'active' (selected), how many are selected (quantity), and whether they are straight, bent, flat, or curved (joints). It is at this level (i.e., the feature) that the minimal units of contrast can be identified. For example, BSL GAY and BSL UNSURE in Figure 16.2 differ according to features of selected fingers: GAY is specified only for the thumb (i.e., no finger features), whilst UNSURE is specified for the thumb and [all] fingers.

16.2.3.2 Place of Articulation (POA)

As with handshape, POA is represented within the Inherent Features structure. Its organisation within the Prosodic Model reflects the generalisation that there are four major body regions (the body, the head, the torso, and the arm) and that each region has eight place distinctions. Beginning from the POA branch within the Inherent Features structure, the passive articulator is specified. This is divided into three-dimensional planes – horizontal (y-plane), vertical (x-plane), and mid-sagittal (z-plane). Signs occurring along the vertical plane may also be specified for one of the major locations on the body: the head, the torso, the arm, or the non-dominant hand. Within each of the eight major locations, eight further POA values are specified. For example, the eight POA values for the head, which are thought to be contrastive in ASL, are: top of the head, forehead, eye, cheek/nose, upper lip, mouth, chin, under the chin. The model predicts that there will be eight distinctions in each of the major locations, but the values may well be language particular, differing from sign language to sign language.

16.2.3.3 Orientation

Orientation is traditionally regarded as a minor parameter since there are fewer minimal pairs based on orientation alone (Brentari, 2012). Earlier descriptions of orientation (e.g., Stokoe et al., 1965; Battison, 1978) were often based on the direction of the palm and the fingertips (e.g., in BSL UNSURE, the palm is facing leftwards and the fingertips are facing forwards). Within the Prosodic Model, as well as Crasborn and van der Kooij (1997) for Sign Language of the Netherlands (NGT), orientation is regarded as being derived from a relationship between an active handpart and the POA. From this perspective, the orientation in BSL UNSURE would be expressed as the relation between the ulnar side of the dominant hand (i.e., handpart) towards the palm of the non-dominant hand (i.e., the POA).

16.2.4 The core lexicon: Prosodic Features

Returning to Figure 16.3, one can see that the root lexeme branches into both Inherent Features and Prosodic Features. Figure 16.5 provides a detailed representation of the organisation of the Prosodic Features tree.

Within the Prosodic Features structure of the Prosodic Model, the dynamic elements of signs are specified. These dynamic elements contrast with handshape and

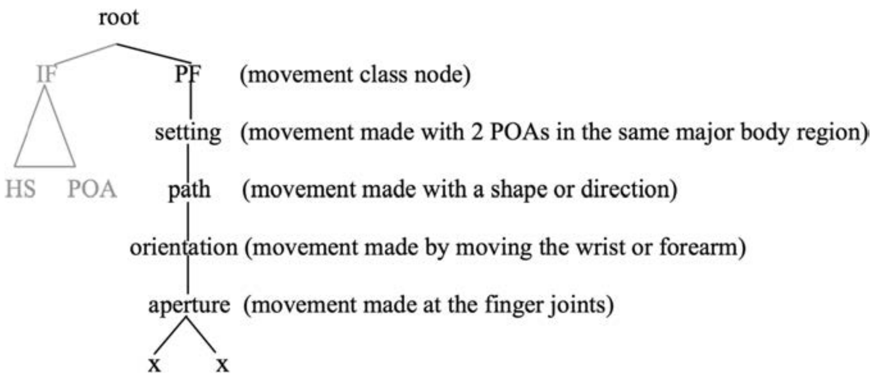


Figure 16.5 Prosodic Features representation within the Prosodic Model

POA within the Inherent Features branch since, by their very nature, they are features that can change within a sign. Additionally, it is within the Prosodic Features branch that segmental (or timing units) structure is derived. A major motivation in this separation lies in the fact that Prosodic Features are realised sequentially, whilst Inherent Features are realised simultaneously. In addition, the hierarchical structure within the Prosodic Features branch is not as complex when compared to the organisation of Inherent Features.

16.2.4.1 Movement

Movements are dynamic acts with a trajectory, a beginning, and an end; their phonological representation will vary depending on the body part used to articulate the movement (see Figure 16.6). The movement change in ASL UNDERSTAND is a change in aperture that is articulated by the finger joints. The movement change in ASL HAPPEN is a change in orientation articulated by the radial-ulnar (forearm) joint. Furthermore, it is the elbow that articulates a path movement in ASL SEE and the shoulder that articulates a setting movement in ASL WE. Body parts involved in the articulation of movement are organised within the Prosodic Model beginning with the more proximal joints (e.g., the shoulder) and ending with the more distal joints (e.g., the fingers). In some signs, it is also possible to have two simultaneous types of movements articulated together. For example, ASL HAPPEN can also be produced with both the radial-ulnar joint and movement from the shoulder joint resulting in a downward movement. Other signs like ASL THROW have both a path movement and what is known as a secondary movement (i.e., aperture change within the hand) (Sandler and Lillo-Martin, 2006). The different types of movements as they would be represented within the Prosodic Model are also provided in Figure 16.6.

Whilst much has been made of the simultaneous nature of sign languages, it is uncontroversial (as noted above) that signs are comprised of sequential elements. This sequentiality is represented through timing slots projected within the prosodic structure (shown as x-slots in Figure 16.6). Path features generate two timing slots; all other features generate one timing slot. Inherent Features do not generate timing slots at all; only movement features do this within the Prosodic Model. When two movement components are articulated simultaneously as in ASL THROW, they align with one another, and only two timing slots are projected onto the timing tier (see ASL THROW in Figure 16.6). Timing slots typically do not create

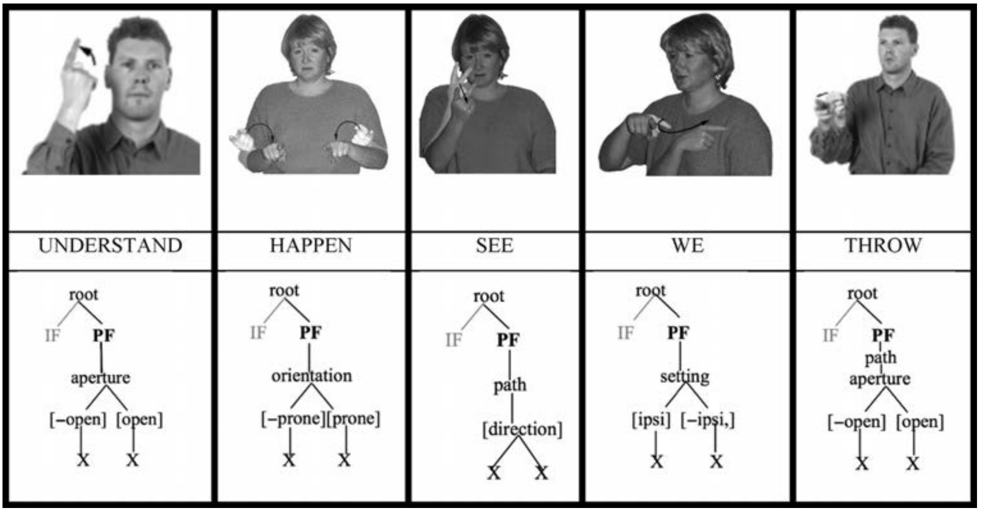


Figure 16.6 Different types of movement in ASL and as represented within the Prosodic Model

minimal pairs (i.e., duration is not contrastive in sign languages) but play an important role in describing where morphological modifications appear. For instance, when a sign is modified for intensity in both ASL and BSL, the first segment is lengthened (e.g., BSL QUICK can be held in its initial position during its articulation resulting in the overall meaning of 'very quick').

16.2.5 Phonological units in sign language

In our brief description of the parameters recognised as playing a role in the phonology of sign languages, one can see that parallels can be made with phonological units attested in spoken language. A *parameter* in sign languages constitutes a fundamental group of *features*, similar to possible segment types in spoken languages (e.g., vowels, glides, obstruents, approximants). A group of features is often referred to as a 'major class' in general phonological theory, specifically in feature geometry – e.g., 'laryngeal' or 'dorsal' are feature classes in spoken languages, and are at the same level as 'joints' or 'selected fingers' within the handshape parameter. Consequently, features such [+flexed] and [−flexed] have the same relation to the 'joints' feature class in a sign language as [spread glottis] has to the 'laryngeal' class in a spoken language. These features, as in spoken language phonology, are the smallest units and the minimal properties that can create a minimal pair.

Movement features also play an important role in the sign language syllable with movement being described as analogous to vowels. Parallels between the two can be seen when one considers that vowels and movements are perceptually the most salient feature within a word or a sign and that movements are what makes signs visible, just as vowels make words audible. In fact, researchers have proposed that more visually salient movements are more sonorous – that is, wiggling the fingers is less sonorant than twisting of the radial-ulnar joint (forearm), which is less sonorous than a path movement (Sandler, 1993; Corina, 1990; Brentari, 1993; Perlmutter, 1992). The criteria for counting syllables in sign languages are outlined in Figure 16.7.

Syllable counting criteria: The number of syllables in a sequence of signs equals the number of sequential movements in that string	
a.	When several shorter (e.g., secondary) movements co-occur with a single (e.g., path) movement of longer duration, the longer movement is the one to which the syllable refers
b.	When two or more movements occur at exactly the same time, it counts as one syllable, e.g., ASL THROW is one syllable containing an aperture change and a path movement

Figure 16.7 Syllable counting criteria
(Brentari, 1998)

Several arguments can be made to demonstrate that movement plays a central organising role at the phonological level forming a unit similar to the syllable nucleus in spoken languages. Firstly, fingerspelled letters or number signs produced in stasis have been observed to add an epenthetic movement in some sign languages when used as an independent word (Brentari, 1990; Jantunen, 2007; Geraci, 2009). Brentari (1990) suggests that, as in spoken languages where an operation of vowel epenthesis ensures syllable well-formedness, movement is inserted where necessary to ensure that the signed output is a well-formed syllable. Secondly, the repetition of movement appears as a rhythmic sequential unit produced by deaf infants at a similar milestone to vocal babbling observed in hearing children (Pettito and Marentette, 1991). Thirdly, morphological modifications to signs are often permitted on the basis of their movement properties. Signs containing one movement element are permitted to undergo modifications that result in derived nominal forms (e.g., the path movement in ASL SIT can be repeated to derive the nominal form CHAIR) in contrast to signs consisting of two or more movements such as ASL THROW (which contains both a path and secondary movement). This suggests that forms allowing reduplication have one simultaneous movement component and are light syllables, whilst those that disallow reduplication have two or more simultaneous movement elements and are therefore heavy. This also demonstrates that sign syllables do not have the same internal structure as spoken languages syllables – i.e., syllable weight and sonority are not related in this way in spoken languages.

Lastly, it should be noted that the parameters specified (i.e., handshape, POA, and movement) all combine to form a lexeme at the root node in contrast to spoken languages where they would combine to form a vowel- or consonant-like unit. As mentioned above, this demonstrates that features in sign languages are typically specified only once per lexeme, not once per segment or once per syllable, but once per word. This is a fact that is – if not explicitly stated – implied in many models of sign language phonology. Whilst parallels can be drawn with tone in tonal languages and features that harmonise across a lexeme (e.g., vowels), it appears that fewer features in speech are associated with the domain of the word in spoken languages than in signed languages; this points to a fundamental difference between signed and spoken language phonology. Importantly, all sign languages that have been subject to serious inquiry have been noted to operate in this way; the extent to which tone and vowel harmony are attested cross-linguistically for spoken languages does not approach a similar scale by comparison.

16.3 Critical issues and topics

In this section, we underline three areas of interest within the field of sign language phonology. These are: the development of phonological models/phonological theory, the effect of modality on phonological organisation, and the link between sign language and gestures produced by non-signers at the phonological level. The third area is one that has received particular interest from sign language linguists in recent years.

16.3.1 Phonological models

Different phonological models have been proposed to account for the underlying representation of signs. These can be understood with reference to the history of phonological theory generally. The earliest model of sign language phonology proposed by Stokoe (1960) emphasised the simultaneous nature of signs (i.e., the parameters of handshape, location, and movement are all realised at the same time in production) and made no attempt at defining these parameters according to a hierarchy. Instead, like spoken language models in the 1950s, Stokoe (1960) focused on providing evidence for the feature units using phonemic contrast (as explained above). Later models made the important observation that signs could also be comprised of sequential segments (or timing units). Beginning with Liddell and Johnson's (1989) Hold-Movement Model, a sign was divided into linear segments described as either 'holds' or 'movements' at the centre of its representation. Within each segment a number of articulatory features could be identified, although these features did not appear to enter into a hierarchical relationship with one another. Liddell and Johnson's model can be said to mirror Chomsky and Halle's (1968) *Sound Pattern of English*, which was biased towards a sequential representation of segments. Liddell and Johnson drew parallels between spoken and signed languages by likening holds (i.e., static elements) to consonants and movements (i.e., dynamic elements) to vowels. However, as features were individually associated within each segment, the Hold-Movement Model contained a substantial amount of redundant information (e.g., for the BSL sign NAME as shown in Figure 16.2, the same handshape would be specified across segments despite there being no change in this parameter). As spoken language models became increasingly non-linear, the Hand Tier Model (Sandler, 1989), the Prosodic Model (Brentari, 1998), and the Dependency Model (van der Kooij, 2002) would unite both the sequential and simultaneous nature of signs in their respective representations. These models used feature geometry to hierarchically organise a sign's parameters according to their phonological behaviour and articulatory properties. The Hand Tier Model would first address the shortcomings of the Hold-Movement Model by representing handshape as an autosegment. Although linear sequential segments continued to occupy a central role in this model, the simultaneous nature of the sign was also acknowledged. In contrast, later models such as the Prosodic Model and the Dependency Model (van der Hulst, 1993) both placed the simultaneous structure back in central position. Although they differ in some details, both models suggested that segmental structure, despite playing an important role in phonology, is derived from the features specified within a sign. Within the Dependency Model, segmental structure is linked to handshape and POA. Movement is given a minor role within the representation since van der Hulst argued that movement could be derived from handshape and POA features. In contrast, the Prosodic Model acknowledged that handshape, POA, and movement all have autosegmental properties but inherent and prosodic elements were placed on separate branches, and it is within the latter branch that segmental structure is derived.

Focusing on recent models that have adopted a featural and autosegmental perspective within their representations (e.g., the Hand Tier Model, the Prosodic Model, and the Dependency Model), we can see that there is much consensus across representations despite their differences. For example, there is a tendency for features within the parameters of handshape and POA to be specified once per lexeme. For handshape, this generalisation is captured by the *Selected Fingers Constraint* (Mandel, 1981; Brentari, 1998) (or the *Handshape Sequence Constraint* in the Hand Tier Model (Sandler, 1989)), which specifies that a sign only has one set of selected fingers within its articulation; note that ASL THROW, despite having a handshape change, has the same number of selected (or extended) fingers at the beginning and end of its articulation. However, there are important differences between these models that pose interesting questions for further research within the field. These differences point to conflicting ideas regarding the underlying role of a particular articulator or aspect within a phonological representation and highlight areas for further research.

One example is the role of the non-dominant hand in phonological models. Although both hands are often active in connected signing, linguists typically refer to one hand as the primary articulator (or the dominant hand) and the other as the passive articulator (or the non-dominant hand), and distinguish between three types of two-handed signs (Battison, 1978). In Type 1 signs both hands share the same handshape and movement – they are symmetrical (e.g., BSL PAPER and BROTHER in Figure 16.2); in Type 2 signs both hands share the same handshape, but not the same movement – one hand is held stationary (e.g., BSL UNSURE in Figure 16.2); and in Type 3 signs the hands share neither the handshape nor the movement – again, one hand is held stationary (e.g., BSL GAY in Figure 16.2). Each phonological model varies in its treatment of the non-dominant hand within its representation. For the Hand Tier Model, the non-dominant hand has a dual role as a POA in Types 2 and 3 signs and as an active articulator in Type 1 signs. In contrast, within the Prosodic Model and the Dependency Model, the two functions are united and the non-dominant hand is represented as having a dependent role (see Figure 16.4 where H2 occupies the dependent branch of the manual node of the Prosodic Model) within their representations. This captures the fact that the degree of complexity on H2 is severely constrained, an observation made by Battison (1978) when he formalised the *Symmetry* and *Dominance Constraints*. That is, the non-dominant hand is either specified for the same handshape and movement as the dominant hand (the Symmetry Constraint), or if the non-dominant hand is stationary, the handshapes we can expect to see on this hand is restricted to a limited set (the Dominance Constraint).

Additionally, the role of non-manual features (e.g., face and body) within phonological models is unclear. Non-manual features are frequently cited in the literature as a significant parameter in addition to parameters involving the hands. Within the Prosodic Model, non-manual features are represented within the Inherent Features branch in the top-most node of the Handshape structure (see Figure 16.4). Signs can be articulated using non-manual features alone and pairs of signs (featuring a manual component) can be contrastive along the non-manual dimension. For example, the ASL signs LATE and NOT-YET differ only in the presence of tongue protrusion in the latter. However, it should be noted that non-manual signs are extremely infrequent when compared to manual signs and very few minimal pairs exist along this dimension. In addition, non-manual features, such as the eyebrows, appear to play an important role at the suprasegmental level and have been likened to intonation in spoken languages (Nespor and Sandler, 1999; Sandler and Lillo-Martin, 2006). These markers also appear to play a role in morphology and syntax (e.g., Neidle et al., 2000; Zeshan, 2004). Given their minor role at the phonological and lexical levels and their other roles in morphology, syntax, and discourse, it is unclear how non-manual features should be represented within phonological models. Indeed, current models proposed for sign languages often lack an adequate representation for this parameter.

Another area where models differ is in the representation of the movement parameter. The Prosodic Model ascribes a central role to movements. The structure in Figure 16.5 captures not only the phonological features of movement, but also provides a coherent backbone for the syllable and foundation for higher order prosodic structure. There is widespread agreement that movement plays a key role in syllable structure with regard to its functional similarity to vowels and to syllable nuclei, as well as in higher order prosodic structure – for example, the phenomenon of phrase-final lengthening in Intonational Phrases. However, some models have tried to avoid representing movement as a major parameter, and instead derive the movement from the locations and orientations of the start and end of the syllable, with some additional features for manner and repetition (Uyechi, 1995; Channon, 2002; Channon and van der Hulst, 2011).

Finally, it should be noted that although minimal pairs can be found for most parameters in ASL and BSL, it is difficult in many cases to identify minimal pairs for every purported phonological value (i.e., every handshape, location, movement, or orientation) that has been argued to be contrastive within these languages. The only exhaustive attempt to do this for any sign language that we know of is Johnston (1989) for Auslan (Australian Sign Language). More evidence is needed about lexical contrast in ASL and BSL before claims about particular contrastive units can be confirmed. In the meantime, it has been proposed that phonetic structures, including features, should be judged to be phonologically relevant if they (i) create a minimal pair, (ii) are involved in a phonological rule, or (iii) are morphological.

16.3.2 Modality

The second issue we discuss here concerns the effect of modality on phonological organisation. The articulators involved in speaking and signing are different; the articulators in speech are the lips, teeth, tongue, throat, and larynx and the articulators in signing are the hands, arms, head, body, and face. As outlined by Meier (2002) and Meier (2012), there are fundamental differences between these sets of articulators. Firstly, the primary articulators involved in sign languages are paired; there are two hands and arms involved in articulation, whilst there is only a single articulator involved in speaking. As phonology is the level of the language that directly interfaces with the articulators, anatomical differences, in turn, have the potential to influence the phonological structure of languages across modalities. It has been proposed that the organisation of a syllable in speech stems from the opening and closing movement of the jaw which acts as an oscillator in speech (MacNeilage, 2008; MacNeilage and Davis, 1993). When one looks at sign languages, it is apparent that there is not a single oscillator linked to articulation. Signs can be produced by different joints within the arms and the hands. On this basis, Meier (2012, 2002) concludes that the syllable in sign language is physically distinct from the syllable in spoken languages since it clearly has a more varied articulatory basis.

The fact that these larger articulators have an effect on production is evidenced by the rate at which words or signs are produced. Studies have reported that the rate of signing appears to be much slower when compared to speaking (Klima and Bellugi, 1979; Grosjean, 1977). In a study by Bellugi and Fischer (1972), the rate of signing – measured as signs per second – was twice as long as the rate of speaking – measured as words per second. This difference in production may be attributed to the size of the articulators as the arms and hands are much larger and therefore require more effort to move than those involved in speaking (e.g., the jaw and the tongue). Despite the slower rate of signing compared to speech, however, Bellugi and Fischer found that the proposition rate was similar across signed and spoken languages. They attributed this to the use of simultaneous organisation in sign languages, concluding that both modalities are equally efficient at conveying information, but do so in different ways.

There are also differences in perception (Brentari, 2002). In audition, humans can temporally resolve auditory stimuli when they are separated by an interval of only 2 milliseconds (Green, 1971; Kohlrausch et al., 1992), whilst the visual system is much slower and requires at least 20 milliseconds to resolve visual stimuli presented sequentially (Chase and Jenner, 1993). The advantage of temporal processing therefore goes to audition. In contrast, simultaneous processing advantages vision over audition. The effect of the speed of light transmission on the perception of objects is that vision can take advantage of light waves reflected from the target object together with secondary reflection from other objects in the environment onto the target object (i.e., visual 'echo' waves). The combination of the two, perceived simultaneously, enhances the three-dimensional quality of the target object (Bregman, 1990) and allows a three-dimensional image to be perceived quickly due to properties of the signal (the same echo phenomenon in audition is much slower). Given these differences in perception across modalities, one might expect words in signed and spoken languages to exploit the advantages available to their respective systems.

One outcome of this basic design of the auditory and visual physiological system is the effect on word shape. Sign languages have a strong tendency towards being monosyllabic. In Stokoe et al. (1965), 83% of the lexical entries are composed of single sequential movements (using the syllable counting criteria in Figure 16.7). Evidence for this tendency towards monosyllabicity can also be seen in compounds and nativised fingerspelled signs (i.e., fingerspelled sequences that move from the non-native lexicon to the core lexicon over time, thus taking on phonological characteristics of signs from the core lexicon). This monosyllabic nature is retained even when signs are meaningfully modified in a number of ways (i.e., these modifications are typically feature-sized and simultaneously layered onto the stem). This patterning between meaningful elements and phonological structure represents a substantial difference between sign languages and spoken languages. Whilst spoken languages do have simultaneous phenomena in phonology and morphophonology such as tone, vowel harmony, nasal harmony, and ablaut marking (e.g., the past preterite in English – sing/sang; ring/rang), this does not approach the scale of simultaneity seen in signed languages. This pattern demonstrates that signal processing differences in the visual and auditory system clearly have an effect on language typology across modalities.

Modality can also have an effect on the distribution of phonological features. In sign languages, the addressee must look at the person signing to them. Since visual acuity is greater toward the central vision area than in the peripheral areas, we might expect an effect on the distribution of features. This appears to be the case for both ASL and BSL, regarding the distribution of marked and unmarked handshapes. In both Battison (1978) for ASL and BSL SignBank (Fenlon et al., 2014) for BSL, when examining signs produced on the body, signs with a marked handshape (i.e., handshapes which are less salient and more difficult to perceive quickly) were much more likely to be produced in the head and neck locations over the trunk and arm locations. For example, in BSL, out of a possible 376 signs using a marked handshape, 286 (76%) are produced on the head and neck locations where visual acuity is greatest). Similarly, one-handed signs (e.g., BSL NAME as in Figure 16.2) are much more likely to occur in the head and neck locations over Type 1 two-handed signs (e.g., BSL PAPER as in Figure 16.2). Additionally, 81.7% (517/633) of BSL signs produced in the head and neck locations are one-handed compared to 59.9% (169/282) of signs produced in the trunk and arm locations. Siple (1978) suggests that, in conditions of lower acuity, more redundancy may be present in the signal. For Type 1 two-handed signs produced on the trunk, having both hands behave in an identical fashion in the periphery of the addressee's vision means there is more information available to the addressee to identify the sign. This observation, together with the distribution of marked and unmarked handshapes with respect to location, suggests that the constraints imposed on the distribution of features have their origins in perception as suggested by Siple (1978) and Battison (1978).

To sum up, one might therefore expect words in signed and spoken languages to exploit the advantages available to their respective systems. As phonology is the level of the grammar that has a direct link with the perceptual and articulatory phonetic systems, whether visual-gestural or auditory-vocal, we might expect to see differences emerge between the two types of languages in their organisation of phonological elements. This allows us to question to what extent we can see phonological patterns that are similar across the two modalities and to what extent they are different. These findings have implications for the understanding of phonological theory in general.

16.3.3 Sign language and gesture

One of the most debated issues in sign language phonology (indeed in sign language linguistics generally) is the relationship between sign language and gesture. Once the field of sign language linguistics began as an area of serious enquiry, the focus was on making sure that sign languages were credited as linguistic systems distinct from gestures produced by hearing non-signers. Prior to this point, work by scholars often presented a misleading and ignorant view of sign languages, considering them to be a primitive form of communication. Given such opinions, together with the exclusion of sign languages from deaf education and their low status in mainstream society, it is not surprising that suggested associations with gesture since that time have been met with resistance from those in the field of sign language research. It was not until the 1990s that researchers began to seriously consider the relationship between sign language and gesture (Emmorey, 1999; Liddell, 2003; Liddell, 1990; Brennan, 1992).

One area in which the issue of gesture has been most prominent is within the literature on classifier constructions in sign languages. These are signs that occupy the non-core native lexicon and are also known as classifier signs, classifier predicates, depicting signs, depicting constructions, or polymorphemic signs. The handshape identifies the class of the referent and under most analyses is considered to have a morphemic status (e.g., Supalla, 2003; Liddell, 2003). Handshapes may represent classes of objects, either partially or wholly, or the handling of objects (handling constructions). Both types of handshapes are provided in Figure 16.8.

Signs from the non-core native component differ from signs within the core lexicon in that they are highly variable and weakly lexicalised. In terms of their phonology, classifier constructions do not always adhere to phonological constraints to the same extent as core native lexical signs (Aronoff et al., 2003; Eccarius and Brentari, 2007). In addition, whilst handshapes that make up the signs within the core lexicon are argued to be purely phonological, handshapes within classifier constructions carry meaning and are considered to be additionally morphemic.

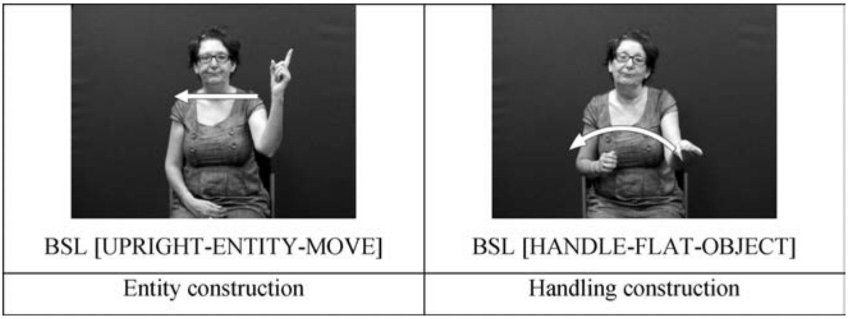


Figure 16.8 Entity construction and handling construction
(pictures from Cormier et al., 2012)

There are two opposing opinions regarding the status of these constructions: some view these constructions and all their components as part of a linguistic system that can be compared to spoken language classifier systems (Emmorey, 2003; Zwitserlood, 2012; Supalla, 2003), whilst some take the view that these constructions are different from spoken language classifier systems and include some gestural elements (Schembri, 2003; Liddell, 2003). Alternative terminologies used to refer to these forms often reflect this opinion, e.g., ‘depicting signs’ or ‘depicting constructions’. One argument supporting the view that classifier constructions in sign languages are unlike spoken language systems is that these forms (specifically entity and handling constructions) have much in common with observer and character viewpoint gestures, respectively (Cormier et al., 2012). This similarity raises an interesting question from a phonological perspective regarding the extent to which constructions within this component demonstrate evidence of phonological patterning not seen in gesture: to what extent do the two differ?

Partly in response to this question, there has been a growth in the number of studies comparing the structure of signs by deaf signers with the gestures produced by non-signers. One such study, using the Verbs of Motion Production task (Supalla et al., n.d.) to elicit entity constructions from Auslan signers and gestures from non-signers without speech, has demonstrated that entity classifier constructions differ from gestures in that signers tend to draw upon a smaller, more conventionalised set of handshapes to represent various entities than non-signers (Schembri et al., 2005). However, less difference was seen between signers’ and non-signers’ use of movement and spatial arrangement of the two hands. In a similar study, Brentari et al. (2012) investigated the use of entity (those that represent the properties of a particular object) and handling constructions (those that represent how one handles a particular object) produced by signers of ASL and Italian Sign Language and compared them to entity and handling gestures produced by Italian and English non-signers in silent gesture mode (i.e., without speech). Participants had to describe what they had seen in vignettes that displayed objects with or without an agent manipulating them. Whilst the two groups of signers and two groups of gesturers patterned similarly to one another, differences were seen between signers and gesturers. Signers displayed more finger complexity in object handshapes, whilst gesturers displayed a tendency to show more finger complexity in handling handshapes (i.e., measured in terms of selected fingers complexity). Brentari et al. suggest that whilst the gesturers attempted to directly imitate the handling handshapes they saw within the vignettes, the signers drew upon the inventory of handshapes available to them within their languages (which were less complex than the real-life action of handling objects in terms of selected fingers complexity). A follow-up study analysing the productions of the use of handling and object handshapes for agentive/non-agentive events by the same groups found that, whilst signers do and gesturers in general do not make this distinction, more Italian gesturers were able to produce this opposition than American gesturers. This suggests that some gesturers, particularly those from cultures that are observed to gesture more frequently, can intuit how to produce this distinction under laboratory conditions (Brentari et al., 2015).

Evidence from psycholinguistic studies have demonstrated that native signers can categorically perceive contrastive handshapes within the core lexicon, whilst gesturers do not (Emmorey et al., 2003). However, there is evidence to suggest that categorical perception of classifier handshapes is not restricted to signers alone. Sevcikova (2013) and Sehyr Sevcikova and Cormier (2016) demonstrate that handling handshapes for flat, rectangular objects (e.g., books) and cylindrical objects (e.g., jars) varying in aperture are perceived categorically by deaf BSL signers and hearing non-signers. Such a finding points away from a phonemic categorical distinction for handling handshapes in the non-core lexicon and towards a more conventionalised gestural system shared by deaf signers and hearing non-signers. An additional follow-up study by Sevcikova (2013) examined whether continuous variation in the size of object was categorically encoded in the production of handling constructions by deaf signers and hearing gesturers (in a co-speech condition and a silent pantomime condition). Participants were first presented with stimuli encouraging them to describe the handling of various objects differing in size and these productions were later matched with the original item by a second group of judges. Within hypothesised categories of graspable object sizes (following Goldin-Meadow et al., 2007), deaf participants judging handling constructions by other deaf signers and hearing participants judging handling constructions with speech were at chance matching items with handshapes for both object types. In contrast, hearing participants judging handling constructions produced during pantomime (gesture without speech) displayed continuous size encoding for both object types.

The results from the experimental studies mentioned above allow us to make several conclusions regarding the nature of entity and handling constructions and their relationship with gesture. The data from Schembri et al. (2005) and Brentari et al. (2012) demonstrate that signers use a more conventionalised set of handshapes than gesturers. Since there are more differences in handshape than in the representation of location and movement by signers and gesturers in these constructions, Schembri et al. (2005) suggest that they are blends of a linguistically specified handshape which fuses gestural elements of location (and possibly movement), thus providing evidence for the (partly) gestural analysis of these constructions. Additionally, regarding handling constructions at least, sign-naïve participants are able to categorically perceive handshape distinctions in handling constructions and also encode and decode conventionalised handshape categories in co-speech gesture (Sevcikova, 2013). This, together with other findings from Brentari et al. (2015), indicates that signs from this component of the lexicon maintain close links with gesture.

16.4 Current contributions and future research

In this section, we focus on current contributions to the field of sign language phonology. These contributions represent new directions within the field as they incorporate new methods and technologies to help us better understand sign language phonology. Looking back at work on sign languages since the 1960s, one can see that early research was aimed primarily at showing that phonology exists in sign languages (e.g., Stokoe, 1960; Klima and Bellugi,

1979) and that units compatible with spoken languages can be identified and organised in a similar way within phonological models (as outlined in section 16.2.5). Whilst these works represent significant advances within the field, there remains much to be uncovered. In this section, we present recent research on understudied sign languages, sign language corpora, and approaches using instrumented capture which have the potential to further illuminate our understanding of the field. The studies referred to here are not intended to be exhaustive but are presented to the reader as an example of how the field can benefit from these directions in future.

16.4.1 Documentation of sign languages

Earlier work on sign language phonology sometimes worked on the assumption that an insight from one sign language was likely true for all other sign languages. Indeed, cross-linguistically, it is easy to see that there are many similarities across sign languages in phonological structure. For example, the Selected Fingers Constraint appears to hold generally for sign languages. However, much of the published research on sign languages has focused on languages based in North America and Northern Europe. Our understanding of phonological structure can benefit from insights gained from other sign languages that have not been well studied to date. For example, work including sign languages in Asia has extended representations within phonological models to account for a wider range of handshapes. As noted above, phonological models to date have made a distinction between *selected fingers* and *unselected fingers* in a given sign. Selected fingers refer to the fingers that appear to be foregrounded and, for signs with handshape change, can change position within a sign (e.g., in BSL AFTERNOON and ASL THROW, the index and the middle finger are the selected fingers). Cross-linguistic research, particularly from Asian sign languages (Eccarius, 2008; Eccarius, 2002; Fischer and Dong, 2010), has revealed that, whilst this distinction is sufficient for capturing the majority of contrasts in sign languages, a further distinction needs to be made along the lines of *primary selected fingers* (SF1) and *secondary selected fingers* (SF2), along with *unselected fingers* (−SF). Some examples of handshapes from Hong Kong Sign Language (HKSL) illustrating this point are provided in Figure 16.9. The distinction

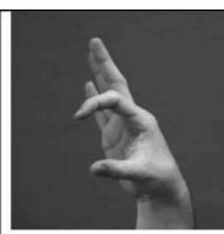
		
DIVIDE	FOX	WC (*Water Closet*)
SF1: **index** SF2:**middle** –SF: **thumb, ring, pinky**	SF1: **thumb, middle, ring** SF2: **index, pinky**	SF1: **thumb, index** SF2: **middle, ring, pinky**

Figure 16.9 Sets of complex handshapes in HKSL displaying primary selected fingers (SF1), secondary selected fingers (SF2), as well as unselected fingers (–SF)

between primary and secondary selected fingers has since been incorporated in recent representations of the Prosodic Model (Eccarius, 2002).

In recent years, more and more work has been conducted on younger sign languages (although it must be said that ASL and BSL can still be described as young languages when compared to spoken languages). This type of research is important for determining how phonology appears in the early stages. In other words, what aspects of a sign are complex at first and become simple over time; what aspects do the reverse? Examples of this type of work can be found in Morgan and Mayberry (2012) where the complexity of two-handed signs is investigated in Kenyan Sign Language with reference to the Symmetry and Dominance Constraint and in Sandler et al. (2005) where the emergence of phonological structure is charted in Al-Sayyid Bedouin Sign Language (ABSL). Both Kenyan Sign Language and ABSL are relatively young, neither more than 75 years old. Data from younger sign languages like this may also contribute to issues of debate within various phonological models. For example, the data from Morgan and Mayberry (2012) appear to support a united view of the dual role of the non-dominant hand in phonological representations (as in the Prosodic and Dependency Models) since exceptional signs are better accounted for when complexity is calculated by taking features of both hands into account each time. In conclusion, it stands to reason that further research on the world's sign languages and their respective lexicons have the potential to extend existing representations so that they may better account for the full range of variability in the world's sign languages.

16.4.2 Studies using sign language corpora

In recent years, there has been a growth in the number of sign language corpus projects world-wide (e.g., in Australia, the Netherlands, Poland, Sweden, and the United Kingdom). Such corpora are designed to be large machine-readable datasets of semi-spontaneous data and representative, as far as is possible, of the deaf community. Once annotated, these corpora can provide us with a unique snapshot of sign language phonology in action across a wide range of social groups, something that has not been possible before. Usage-based studies utilising large datasets including corpora have already investigated some linguistic and social factors that condition phonological variation in sign languages. For example, two sociolinguistic studies, one focusing on BSL (Fenlon et al., 2013) and another on ASL (Bayley et al., 2002), investigated variation in the 1-handshape, a very common handshape used in lexical signs as well as pointing signs such as pronouns, and found that handshape assimilation was conditioned by grammatical category (lexical signs were more likely to preserve handshape; pronominal signs were least likely to do so) and the immediate phonological environment (preceding signs with different handshapes were more likely to assimilate to the target sign). The BSL study reported that the immediate phonological environment was the strongest factor conditioning variation, whilst the ASL study reported grammatical category as the strongest predictor of handshape variation. Similar insights have been made regarding location variation in signs produced on the forehead (e.g., signs similar to BSL NAME) in ASL (Lucas et al., 2002) and Auslan and New Zealand Sign Language (Schembri et al., 2009). These variation studies not only indicate language-specific differences at play but that, in the case of the Auslan study, this variation may be indicative of a change in progress (e.g., the lowering of signs produced on the forehead appears to represent language change in progress led by younger female signers from urban centres). Such studies based on large datasets therefore afford us a strong empirical basis from which observations on phonological change can be made.

Work on sign language corpora has also highlighted the need for lexical databases that are representative of the sign language lexicon since such resources are required to assist with the process of annotation. Databases that have emerged from such projects may include a phonological description of a sign's citation form with each entry (e.g., BSL SignBank (Fenlon et al., 2014)). What this means is, as these resources grow over time to become representative of a sign language's lexicon, we will be in a position to make more accurate generalisations regarding the distribution of phonological properties in these sign languages (e.g., regarding the frequency of handshapes). Previously, even for well-researched languages, such resources have not been available. The association with a sign's occurrence in a corpus of semi-spontaneous data also means that researchers will be able to quickly retrieve tokens of a specific sign to better understand its use and representation in connected signing.

16.4.3 Sign language and phonetic instrumentation

As with spoken languages, a thorough understanding of phonology relies on understanding phonetics. The techniques of phonetic instrumentation with sign languages are one area in which there are still considerable advances to be made. Currently, there is no spectrographic analysis for sign language and most phonetic transcriptions are still done manually with the naked eye (e.g., many researchers use programs like ELAN (www.tla.mpi.nl/tools/tla-tools/elan/) to code data manually). As a result, the field of sign language phonetics is often described as an area in which little progress has been made. In recent years, however, there has been a growth in the number of studies using motion capture in order to understand exactly how the production of signs vary in terms of its kinetic properties. Earlier work incorporating this type of technique can be traced to the 1990s. For example, one study by Wilbur and Zelaznik (1997) investigated how duration and velocity are distributed with respect to prosodic phrasing using a three-dimensional motion analyser system (WATSMART) and found that duration marks final phrase position and that velocity marks prominence.

Other studies which have used motion capture include Cheek (2001) to investigate pinky extension in handshape, Malaia and Wilbur (2011) in their study of ASL verbs with regards to telicity, and Mauk and Tyrone (2012) in investigating location variation in ASL signs produced at the forehead. These studies provide us with further insights into the factors that affect a sign's production using a more fine-grained analysis and in some cases building on conclusions from studies using corpora. For example, whilst the immediate phonological environment has been identified as an important predictor in handshape assimilation in the corpus-based studies above, Cheek (2001), investigating handshape variation, found an effect of phonetic environment which was rate dependent; such a finding suggests that coarticulation effects (rather than assimilation) are at play. Similarly, Tyrone and Mauk (2010), using motion capture data, found signing rate to be a significant predictor of sign lowering, although this was not uniform for all signs; for some signs, there was an interaction with the phonetic environment. In Mauk and Tyrone (2012), a more detailed analysis (again focusing on signs produced on the forehead) took into account the phonetic location of the preceding and following signs and found that the direction of coarticulation was strongest with the following sign (i.e., as the location value of the following sign moved up the vertical axis, so too did the location value of the target sign ASL KNOW). Additionally, the authors demonstrate that it is not only the hands that move but that passive articulators, the forehead in this case, can play a role. In other words, when producing ASL KNOW, the forehead may move to meet the hand; this type of movement was observed at slower signing rates in ASL. Furthermore, Mauk and Tyrone (2012) found that native signers appear to be more compact in their use of signing space when compared to non-native signers which, in turn, may require them to make more subtle phonological contrasts with respect to the location parameter. Importantly, such studies demonstrate that instrumented capture affords us a better insight regarding the complex array of factors that characterise variation than categorical measures (in particular, coding systems that rely solely on the naked eye).

16.5 Conclusion

To conclude, we have described some of the most important areas of research in sign language phonology both historically and currently. At the end of a chapter such as this one, we would like to offer two reasons why phonologists who conduct research on spoken languages should care about the phonology of sign languages. The first has to do with work concerning sign language and gesture, some of which was discussed in section 16.3.3, which is becoming increasingly important in understanding human language. Spoken languages clearly include both speech and the gestures that accompany them as co-speech gestures. There are numerous researchers that have made this claim in psychology, most notably Goldin-Meadow (2003) and McNeill (1992), and they have described many ways that gestures contribute to the meaning of our utterances in a variety of ways. Pinpointing work that is relevant for a volume on phonology, work by Krahmer and Swerts (2007) have opened up a field of inquiry describing the ways that beat gestures, which co-occur with the prominent syllable of a phonological phrase, can influence how we perceive the *vocal* prominence of the prominent syllable. Assuming this body of work continues to gain support, analyses on spoken languages, particularly on spoken language prosody, will routinely include properties of the face and body and will be bi-modal. The insights from work on sign languages where the body is a site for phonological operations such as the syllable will potentially be of great use in that work. The sign language syllable, in particular, offers tools for how to think about the componentiality of gesture in this new area of inquiry that couples gesture with speech in considering the totality of spoken language.

The second reason why phonologists who study spoken languages should be concerned with sign language phonology constitutes the ultimate goal of our work as a whole: namely, to describe the full range of extant phonological systems and to strive to construct theories that can handle both. Hale and Reiss (2008) have gone so far as to propose that the work of describing 'substance-free' phonology is the primary task of phonologists. Although we are still a very long way from achieving that goal, nonetheless we are optimistic. As Stephen Anderson (1989: 803) wrote, "[Phonology] is a domain of human cognition where we probably know more in detail about the specific principles operative within a particular cognitive subsystem than anywhere else, and about the specific representations that play a part in such knowledge." Sign languages are much more than a couple of extra data points on the landscape of possible phonological systems, or a new quirky set of facts that stretch current theory. They are a set of languages with long histories which have generated solutions to building efficient and effective phonological systems with some materials that are the same as those of speech (the same mind/brain) and some that are different. It is the resilient creativity in response to our human need to communicate that gave rise to the range of phonological structures in sign languages. Working on how signed and spoken languages can genuinely be handled by the same phonological tools gets us ever closer to understanding phonology, generally speaking, and for this reason this chapter is written for everyone in the field.

16.6 Further reading

Battison, R. 1978. *Lexical Borrowing in American Sign Language*, Silver Spring, Linstok Press. This text delivers a good overview of the subject area as well as a description of phonological processes observed in fingerspelled loan signs.

Brentari, D. 1998. *A Prosodic Model of Sign Language Phonology*, Cambridge, MA, MIT Press. We recommend this text for those seeking a detailed description of the Prosodic Model as well as background on sign language phonology.

Liddell, S. K. & R. E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies*, 64, 195–278. This paper introduces the Hold-Movement Model mentioned in section 16.3.1.

Sandler, W. & D. Lillo-Martin, D. 2006. *Sign Language and Linguistic Universals*, Cambridge, Cambridge University Press. This textbook provides a detailed overview of sign language phonology across several chapters focusing on the Hand-Tier Model.

Stokoe, W. 1960. Sign language structure: An outline of the visual communication system of the American Deaf. *Studies in Linguistics Occasional Paper 8*. University of Buffalo. Stokoe's seminal paper remains a must for anyone interested in sign language phonology.

16.7 Acknowledgements

This work was supported by funding from the Economic and Social Research Council of Great Britain (Grant RES-620–28–0002), from the Deafness, Cognition and Language Research Centre (DCAL), from the National Science Foundation (BCS-1227908 to D.B.), from the Andrew W. Mellon Foundation, and from the Neubauer Collegium for Culture and Society at the University of Chicago. We thank Sandra Smith for acting as a model in Figure 16.8.

References

Anderson, J. & C. J. Ewen. 1987. *Principles of Dependency Phonology*, Cambridge, Cambridge University Press.

Anderson, S. 1989. The computer and the mind: An introduction to cognitive science by Philip N. Johnson-Laird [Review]. *Language*, 65, 800–811.

Aronoff, M., I. Meir, C. Padden & W. Sandler. 2003. Classifier constructions and morphology in two sign languages. *In:* K. Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Languages* (pp. 53–84). Mahwah, NJ: Lawerence Erlbaum Associates.

Battison, R. 1978. *Lexical Borrowing in American Sign Language*, Silver Spring, MD, Linstock Press.

Bayley, R., C. Lucas & M. Rose. 2002. Phonological variation in American Sign Language: The case of 1 handshape. *Language Variation and Change*, 14, 19–53.

Bellugi, U. & S. Fischer. 1972. A comparison of sign language and spoken language. *Cognition*, 1, 173–200.

Bloomfield, L. 1933. *Language*, New York, Holt.

Bregman, A. S. 1990. *Auditory Scene Analysis: The Perceptual Organisation of Sound*, Cambridge, MA, MIT Press.

Brennan, M. 1992. The visual world of British Sign Language: An introduction. *In:* D. Brien (ed.) *Dictionary of British Sign Language/English*. London: Faber and Faber.

Brentari, D. 1990. Licensing in ASL handshape change. *In:* C. Lucas (ed.) *Sign Language Research: Theoretical Issues* (pp. 57–68). Washington, DC: Gallaudet University Press.

Brentari, D. 1993. Establishing a sonority hierarchy in American Sign Language: The use of simultaneous structure in phonology. *Phonology*, 10, 281–306.

Brentari, D. 1998. *A Prosodic Model of Sign Language Phonology*, Cambridge, MA, MIT Press.

Brentari, D. 2002. Modality differences in sign langauge and phonology and morphophonemics. *In:* R. P. Meier, D. Quinto-Pozos & K. Cormier (eds.) *Modality and Structure in Signed and Spoken Languages* (pp. 35–64). Cambridge: Cambridge University Press.

Brentari, D. 2012. Phonology. *In:* R. Pfau, M. Steinbach & B. Woll (eds.) *Sign Language: An International Handbook* (pp. 21–54). Berlin, Germany: Mouton de Gruyter.

Brentari, D., M. Coppola, L. Mazzoni & S. Goldin-Meadow. 2012. When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Natural Language and Linguistic Theory*, 30, 1–31.

Brentari, D., A. Di Renzo, J. Keane & V. Volterra. 2015. Cognitive, cultural, and linguistic sources of a handshape distinction expressing agentivity. *Topics*, 7, 1–29.

Brentari, D. & C. Padden. 2001. Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. *In:* D. Brentari (ed.) *Foreign Vocabulary in Sign Languages* (pp. 87–119). Mahwah, NJ: Lawrence Erlbaum Associates.

Channon, R. 2002. Beads on a string? Representations of repetition in spoken and signed languages. *In:* R. P. Meier, D. Quinto-Pozos & K. Cormier (eds.) *Modality and Structure in Signed and Spoken Languages* (pp. 65–87). Cambridge: Cambridge University Press.

Channon, R. & H. van der Hulst. 2011. Are dynamic features required in signs? *In:* R. Channon & H. van der Hulst (eds.) *Formational Units in Sign Languages* (pp. 229–269). Berlin: Mouton de Gruyter.

Chase, C. & A. R. Jenner. 1993. Magnocellular visual deficits affect temporal processing of dyslexics. *Annals of the New York Academy of Sciences*, 682, 326–329.

Cheek, A. D. 2001. *The Phonetics and Phonology of Handshape in American Sign Language*. PhD, University of Texas at Austin.

Chomsky, N. & M. Halle. 1968. *The Sound Pattern of English*, New York, Harper and Row.

Corina, D. 1990. Reassessing the role of sonority in syllable structure: Evidence from a visual-gestural language. *Papers from the 26th Annual meeting of the Chicago Linguistic Society: Vol 2: Parasession on the syllable in phonetics and phonology*. Chicago, IL: Chicago Linguistic Society, University of Chicago.

Cormier, K., D. Quinto-Pozos, Z. Sevcikova & A. Schembri. 2012. Lexicalisation and de-lexicalisation processes in sign languages: Comparing depicting constructions and viewpoint gestures. *Lang Commun*, 32, 329–348.

Crasborn, O. & E. van der Kooij. 1997. Relative orientation in sign language phonology. *In:* J. Coerts & H. D. Hoop. (eds.) *Linguistics in the Netherlands 1997* (pp. 37–48). Amsterdam: John Benjamins Publishing Company.

Eccarius, P. 2002. *Finding Common Ground: A Comparison of Handshape across Multiple Sign Languages*, Cambridge, MA, Purdue University.

Eccarius, P. 2008. *A Constraint-Based Account of Handshape Contrast in Sign Languages*. PhD, Purdue University.

Eccarius, P. & D. Brentari. 2007. Symmetry and dominance: A cross-linguistic study of signs and classifier constructions. *Lingua*, 117, 1169–1201.

Emmorey, K. 1999. Do signers gesture? *In:* L. S. Messing & R. Campbell (eds.) *Gesture, Speech, and Sign* (pp. 133–158). Oxford: Oxford University Press.

Emmorey, K. 2003. *Perspectives on Classifier Constructions*, Mahwah, NJ, Lawrence Erlbaum Associates.

Emmorey, K., S. McCullough & D. Brentari. 2003. Categorical perception in American Sign Language. *Language and Cognitive Processes*, 18, 21–45.

Fenlon, J., K. Cormier, R. Rentelis, A. Schembri, K. Rowley, R. Adam & B. Woll. 2014. *BSL Sign Bank: A Lexical Database of British Sign Language* (First Edition), London, Deafness, Cognition and Language Ressearch Centre, University College London.

Fenlon, J., A. Schembri, R. Rentelis & K. Cormier. 2013. Variation in handshape and orientation in British Sign Language: The case of the '1' hand configuration. *Language and Communication*, 33, 69–91.

Fischer, S. & Q. Dong. 2010. Variation in East Asian Sign Language structures. *In:* D. Brentari (ed.) *Sign Languages: A Cambridge Language Survey* (pp. 499–518). Cambridge: Cambridge University Press.

Geraci, C. 2009. Epenthesis in Italian Sign Language. *Sign Language & Linguistics*, 12, 3–51.

Goldin-Meadow, S. 2003. *Hearing Gesture: How Our Hands Help Us Think*, Cambridge, MA, Harvard University Press.

Goldin-Meadow, S., C. Mylander & A. Franklin. 2007. How children make language out of gesture: Morphological structure in gesture systems developed by American and Chinese deaf children. *Cognitive Psychology*, 55, 87–135.

Green, D. 1971. Temporal auditory acuity. *Pschological Review*, 78, 540–551.

Grosjean, F. 1977. The perception of rate in spoken language and sign languages. *Journal of Psycholinguistic Research*, 22, 408–413.

Hale, M. & C. Reiss. 2008. *The Phonological Enterprise*, Oxford, Oxford University Press.

Jantunen, T. 2007. Tavu Suomalaisessa Viittomakielessä [The syllable in Finnish Sign Language; with English abstract]. *Puhe ja kieli*, 27, 109–126.

Johnston, T. 1989. *The Sign Language of the Australian Deaf Community*. Doctoral dissertation, University of Sydney.

Johnston, T. & A. Schembri. 1999. On defining lexeme in a signed language. *Sign Language & Linguistics*, 2, 115–185.

Klima, E. & U. Bellugi. 1979. *The Signs of Language*, Cambridge, MA, Harvard University Press.

Kohlrausch, A., D. Püschel & H. Alphei. 1992. Temporal resolution and modulation analysis in models of the auditory system. *In:* M. E. H. Schouten (ed.) *The Auditory Processing of Speech: From Sounds to Words* (pp. 85–98). Berlin: Mouton de Gruyter.

Krahmer, E. & M. Swerts. 2007. The effect of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57, 396–414.

Liddell, S. K. 1990. Four functions of a locus: Reexamining the structure of space in ASL. *In:* C. Lucas (ed.) *Sign Language Research: Theoretical Issues* (pp. 176–198). Washington, DC: Gallaudet University Press.

Liddell, S. K. 2003. *Grammar, Gesture and Meaning in American Sign Language*, Cambridge, Cambridge University Press.

Liddell, S. K. & R. E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies*, 64, 195–278.

Lucas, C., R. Bayley, M. Rose & A. Wulf. 2002. Location variation in American Sign Language. *Sign Language Studies*, 2, 407–440.

MacNeilage, P. F. 2008. *The Origin of Speech*, Oxford, Oxford University Press.

MacNeilage, P. F. & Davis, B. L. 1993. Motor explanations of babbling and early speech patterns. *In:* B. Boysson-Bardies, S. D. Schonen, P. Jusczyk, P. F. MacNeilage & J. Morton (eds.) *Development Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 341–352). Dordrecht: Kluwer.

Malaia, E. & R. Wilbur. 2011. Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech*, 55, 407–421.

Mandel, M. 1981. *Phonotactics and Morphophonology in American Sign Language*, Berkeley, CA, University of California.

Mauk, C. E. & M. E. Tyrone. 2012. Location in ASL: Insights from phonetic variation. *Sign Language & Linguistics*, 15, 128–146.

McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*, Chicago, IL, University of Chicago Press.

Meier, R. P. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. *In:* R. P. Meier, K. Cormier & D. Quinto-Pozos (eds.) *Modality and Structure in Signed and Spoken Languages* (pp. 1–26). Cambridge: Cambridge University Press.

Meier, R. P. 2012. Language and modality. *In:* R. Pfau, M. Steinbach & B. Woll (eds.) *Sign Language: An International Handbook* (pp. 574–601). Berlin, Germany: Mouton de Gruyter.

Morgan, H. & R. Mayberry. 2012. Complexity in two-handed signs in Kenyan Sign Language. *Sign Language & Linguistics*, 15, 147–174.

Neidle, C., J. Kegl, D. MacLaughlin, B. Bahan & R. Lee. 2000. *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*, Cambridge, MA, MIT Press.

Nespor, M. & W. Sandler. 1999. Prosody in Israeli Sign Language. *Language and Speech*, 42, 143–176.

Perlmutter, D. M. 1992. Sonority and syllable structure in American Sign Language. *Linguistic Inquiry*, 23, 407–442.

Pettito, L. & P. Marentette. 1991. Babbling in the manual mode: Evidence for the ontogeny of language. *Science*, 25, 1493–1496.

Sandler, W. 1989. *Phonological Representation of the Sign: Linearity and Non-Linearity in American Sign Language*, Dordrecht, Foris.

Sandler, W. 1993. A sonority cycle in American Sign Language. *Phonology*, 10, 209–241.

Sandler, W. & D. Lillo-Martin. 2006. *Sign Language and Linguistic Universals*, Cambridge, Cambridge University Press.

Sandler, W., I. Meir, C. Padden & M. Aronoff. 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the Unites States of America*, 102, 2661–2665.

Schembri, A. 2003. Rethinking "classifiers" in signed langauges. *In:* K. Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Languages* (pp. 3–34). Mahwah, NJ: Lawrence Erlbaum Associates.

Schembri, A., C. Jones & D. Burnham. 2005. Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *Journal of Deaf Studies and Education*, 10, 272–290.

Schembri, A., D. McKee, R. McKee, S. Pivac, T. Johnston & D. Goswell. 2009. Phonological variation and change in Australian and New Zealand Sign Languages: The location variable. *Language Variation and Change*, 21, 193.

Sehyr Sevcikova, Z. & K. Cormier. 2016. Perceptual categorization of handling handshapes in British Sign Language. *Language and Cognition*, 8, 501–532.

Sevcikova, Z. 2013. *Categorical versus Gradient Properties of Handling Handshapes in British Sign Language (BSL)*. PhD, University College London.

Siple, P. 1978. Visual constraints for sign language communication. *Sign Language Studies*, 19, 95–110.

Stokoe, W. 1960. Sign language structure: An outline of the visual communication system of the American Deaf. *Studies in Linguistics Occasional Paper 8*. University of Buffalo.

Stokoe, W., D. C. Casterline & C. G. Croneberg. 1965. *A Dictionary of American Sign Language on Linguistic Principles*, Silver Spring, MD: Linstok Press.

Supalla, T. 2003. Revisiting visual analogy in ASL classifier predicates. *In:* K. Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Languages*. Mahwah, NJ: Psychology Press.

Supalla, T., E. L. Newport, J. Singleton, S. Supalla, D. Metlay & G. Coulter. n.d. *The test battery for American Sign Language morphology and syntax*, New York, University of Rochester.

Tyrone, M. E. & C. E. Mauk. 2010. Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics*, 38, 317–328.

Uyechi, L. 1995. *The Geometry of Visual Phonology*. PhD, Stanford University.

van der Hulst, H. 1993. Units in the analysis of signs. *Phonology*, 10, 209–241.

van der Kooij, E. 2002. *Phonological Categories in Sign Language of the Netherlands: The Role of Phonetic Implementation and Iconicity*. PhD dissertation, Leiden University.

Wilbur, R. & H. N. Zelaznik. 1997. *Kinematic correlates of stress and position in ASL*, Chicago, IL, Linguistic Society of America.

Zeshan, U. 2004. Interrogative constructions in signed languages: Crosslinguistic perspectives. *Language*, 80, 7–39.

Zwitserlood, I. 2012. Classifiers. *In:* R. Pfau, M. Steinbach & B. Woll (eds.) *Sign Language: An International Handbook* (pp. 158–186). Berlin: Mouton de Gruyter.

17

Phonology as an emergent system

Diana Archangeli and Douglas Pulleyblank

17.1 Introduction

Linguistic thought since the late 1950s has been heavily influenced by the innateness hypothesis, that "there is a rich, innate, species-specific component of the mind dedicated to language" (Boeckx & Piattelli-Palmarini 2005, 448), known as Universal Grammar (Chomsky 1965), which provides a priori knowledge in order for language acquisition to succeed. Emergent Grammar challenges this approach in terms of the balance between linguistic and non-linguistic factors: Emergence proposes that much of grammar is acquired without making use of the a priori knowledge given by an innate human language faculty (HLF). Rather, a (phonological) grammar is the result of making sense out of the surrounding language primarily through general human cognition (Hopper 1998; Bybee 1999, 2010). The residue, those phenomena that cannot be resolved in this fashion, are the domain of the HLF. It should be stressed that this is not a new view. If we think of a line drawn between zero (no language-specific effects) and 100 (nothing is learned), proponents of Universal Grammar (UG) point to significant language-specific effects and the boundary moving well up the scale; in the work we are reporting on here, our goal is to see just how low the boundary can be set.

In this chapter, we lay out the conceptual underpinnings of Emergent Grammar, provide an overview of what an Emergent Grammar might look like, and explore implications of the model for future research. Our focus is on phonological systems; phonological systems do not occur in isolation, and so we address some aspects of the interface with phonetics and with morphology. We largely neglect other relevant interfaces (e.g. syntax, semantics, lexical access, etc.).

There are several contrasts between the Emergent model and models which assume a rich HLF; we address five differences here.

(i) *The Mapping Problem* is concerned with the task of mapping linguistic categories onto patterns observed in raw data when learning a language. With a rich HLF, the learner must identify patterns, identify the relevant universals, and map the patterns to those universals in order to create a grammar. Under Emergence, once patterns are identified, the grammar is created. The Mapping Problem does not exist. This is a challenge for any model that assumes innate linguistic categories.

(ii) *The Exceptions Problem* arises from the observation that languages are rife with exceptions to the rule – forms that do not follow general patterns. With a rich HLF, the expectation is that there are no exceptions so each exception constitutes a counterexample. Under Emergence, as shown below, generalisations are continually built over (sub)parts of the lexicon. A part can be as small as one item; it can be as large as the entire lexicon. Consequently, exceptions – both individual and classes of exceptions – are both expected and highly informative about the structure of a language. Note that this is in contrast to both standard generative phonology and Optimality Theory, where exceptions are not predicted by the models.

(iii) *The Default Language Problem* addresses the question of why languages have (morpho)phonological rules. The Emergence view takes the position that human minds are particularly adept at finding and canonising patterns, even where none need exist (Gerken & Bollt 2008; Archangeli et al. 2011), predicting grammars with rules and alternations, contra SPE (Chomsky & Halle 1968). Complexity – or divergence from the default language – has been measured in various ways: by the number of rules, the number of elements in a rule, the way constraints are ranked. For SPE, the default language would be the language with the simplest rules – that is, no rules at all, in contrast to the expectation under Emergence. In this way, Emergence is somewhat similar to Optimality Theory (Prince & Smolensky 1993; McCarthy & Prince 1993b). In Optimality Theory, the simplest start state for a grammar in terms of learnability would be one where all Markedness outranks all Faithfulness (Hayes 2004). This grammar would produce too few distinct forms to be feasible, so Markedness constraints would be gradually demoted as a language's forms are encountered – necessarily violating Markedness (Tesar & Smolensky 1998). Hence the "default" start state would be maximally unmarked, but actual grammars would necessarily involve more complexity.

(iv) *The Acquisition Problem* is that children acquire the details of the language they are exposed to – for instance, "high" vowels have different F1 and F2 values depending on the language, VOT varies for aspirated/unaspirated stops depending on the language, even the very modality of communication varies depending on the language a learner encounters – oral or signed. While it is possible to posit that the HLF allows such variation, a much more convincing argument for innateness would be consistency across languages in the face of the physical variety that humans are capable of. We see consistency in broad strokes, but not in the details – consistent behaviour is predicted under Emergence where consistency arises from similar anatomies and similar learning mechanisms, but identity across languages is not expected.

(v) *The Null Hypothesis Problem* raises the point that a model without a rich HLF is a reasonable null hypothesis. It is important to note, however, that the null hypothesis need not be "anything-is-possible." Rather than a priori assuming a rich HLF, the Emergence hypothesis examines language structure from the point of view of normal human cognitive abilities that exist independent of language. The residue, the facts that cannot be accounted for by general cognition, would be the domain of the HLF.

17.2 Critical issues: morphology and phonology in a grammar that is emerging

The Emergence model assumes that a great deal (if not all) of a language is acquired making use of general cognitive strategies (Bybee 1999, 2001, 2010). We assume at least those components listed in (1), each of which evolves as the child matures.

(1) General cognitive tools
  a. Memory
  b. Similarity
  c. Frequency
  d. Symbolic systems
  e. Entropy

We illustrate these briefly, using the example of learning the sounds of a language. First, acquiring a language – even its sounds – requires memory (1a). The sorts of memories involved in language are demonstrably rich (Bybee 2010). As "tokens" or "exemplars" of a language's sounds are laid down in memory, categorisation occurs, for example, creating sound categories. Such sound units are joined into larger units, a process of chunking, or "learning from experience" (Newell 1990, 185), where:

> [a] chunk is a unit of memory organization, formed by bringing together a set of already formed chunks in memory and welding them together into a larger unit. Chunking implies the ability to build up such structures recursively, thus leading to a hierarchical organization of memory.
>
> *(Newell 1990, 7)*

There can be no question that a central part of learning the sounds and morphemes of a language involves committing encountered strings to memory.

However, it is not enough to simply remember the sounds of an utterance – it is critical that the learner identify similarities among sounds and among sound chunks (1b). This allows the learner to recognise that two memories of sound are "the same" despite being physically different, and so begins a classification of language in memory (Goldstone 1994). Note there are related but distinct types of similarity: similarity of the whole and similarity of parts. On the one hand, the learner must be able to ascertain that distinct utterances involve the "same" element. Different productions of a word like [dɑg] *dog*, for example, will exhibit different physical properties, and yet the learner comes to identify multiple utterances of [dɑg] as tokens of the same item. In addition, the learner must learn to identify similar properties of items that are clearly not globally the same. For example, a word like [dɑg] *dog* is similar in ways that need to be defined to words like both [dɑt] *dot* and [dɑgi] *doggie*.

Research in child language acquisition shows that children pay close attention to frequency (1c; Maye et al. 2002; Newport & Aslin 2004); frequency is important in sorting out snippets of the sound stream into sounds (chunks with high frequency) and residue (chunks with low frequency, e.g. snippets corresponding to transitions from one sound to another); the same sorting over larger chunks serves to identify related morphs. Similarly, asymmetries in distributional frequencies lead to generalising patterns. Evidence from covert articulatory regularities (Archangeli et al. 2011; Mielke et al. 2010; Mielke et al. 2016) shows that humans devise language patterns even when there is no overt evidence to do so, which supports evidence in the language acquisition literature for facile and rapid pattern identification: Gerken & Bollt (2008) demonstrates that three distinct exemplars of a language pattern are sufficient for 9-month-olds to formulate a generalisation.

As Deacon (1997) argues, a unique human attribute is our ability to represent our knowledge with a symbolic structure (1d). Within language, we assume that this structure encodes representations of specific utterances (including voice, location, emotional content, etc., as well as linguistic information) as instances of more general entities (words, morphs, sounds), each of which may have further relations to other entities; formalising the grammar involves characterising the nature of these entities and of the relations among them. Learning patterns, we assume, involves structuring representations symbolically.

Finally, we assume that the process of acquiring a language involves a general reduction in the entropy of the postulated system (1e; Goldsmith 2000; Hayes & Wilson 2008; Hall 2009). When a learner begins to generate a language, entropy is large. Very little can be predicted – there is a great deal of uncertainty in the system. As the lexicon and grammar evolve, patterns are identified, structures become regularised and predictable, and the degree of uncertainty in the system decreases.

A growing amount of recent research builds on notions of memory, similarity, and frequency, for example, exemplar models (Lacerda 1998; Pierrehumbert 2001) and (exemplar-inspired) usage-based models (Bybee 2001, 2010). Our proposal – that there is a critical role for building a symbolic grammar by generalising and then generalising over the generalisations in order to maximise entropy – results in a more abstract grammar than is often associated with usage-based and exemplar models. It is unclear to what extent this is a significant difference in the models and to what extent it is a difference in the types of phonological patterns being considered.

17.2.1 How a grammar develops

Consider as a first illustration of how a grammar develops the case of Chichewa vowels, traditionally represented as [i, e, a, o, u] (Harris 1994). Through similarity and frequency, the learner can identify five distinct units, { i, e, a, o, u }. But there are many dimensions of similarity, and thus similarity gives several ways to classify these sounds: these different classifications can give rise to natural classes without recourse to innate distinctive features: (i) the vowels can be grouped together as one "vocalic" set, based on acoustic properties and on where they occur in sound sequences; (ii) they can be grouped by tongue height/F1, { i, u } vs. all others, { a } vs. all others, { i, u } vs. { a } vs. { e, o }; (iii) they can be grouped by articulatory properties, like lip rounding, and effect on F2 – { u, o } vs. all others; (iv) they can also be grouped by alternations; in Chichewa, the lexical entry for the causative is a morph set containing two morphs, { -its, -ets }, as is the applied, { -il, -el }, with [-ets, -el] following { e, o } and [-its, -il] occurring after the other vowels, { i, u, a }.

(2) Chichewa affix morphs

		causative	applied	gloss
a.	{ e, o }	lemb-ets-a	lemb-el-a	"write"
		konz-ets-a	konz-el-a	"correct"
b.	{ i, u, a }	pind-its-a	pind-il-a	"bend"
		put-its-a	put-il-a	"provoke"
		bal-its-a	bal-il-a	"give birth"

There is discussion in the literature about whether the relevant feature for this pattern is a tongue body or a tongue root feature: under Emergence, this is not the question to ask. What is relevant is that the vowels { e, o } form a class apart from { i, u, a }. This partition emerges from the behaviour of these vowels if from nothing else. Importantly, convergence of acoustic, articulatory, and/or behavioural properties (points (i) to (iv)) will make certain classes more robustly identifiable for the learner. Thus, it is no surprise that natural classes in phonological patterns are typically defined in both articulatory and acoustic terms, and that there is robust evidence for these classes from phonological patterns – such cases provide rich evidence for the relevant sound classes. Importantly, however, under Emergence there is no a priori requirement that phonological classes be physically defined and vice versa: there is no requirement that a physical distinction give rise to phonological classes.

Once meanings are associated with sounds, the same bottom-up strategies result in identifying morphological structure and morphotactic relations. Similarity of meaning, function, and sound converge to identify individual morphs and the relations between them (e.g. in North American English, coronal-final words have a [ɾ]-final morph, { weɪt, weɪɾ } "wait"; { weɪd, weɪɾ } "wade," etc.). Productive relations between morphs mean that the learner can effortlessly "fill in the paradigm": there is no need to encounter each morph independently (Archangeli & Pulleyblank 2012). This results in a network of morphs interrelated by sound, by meaning, and by function. This concept of related morphs replaces the more familiar concept of single underlying representations for morphemes along with allomorphy rules/constraints determining the various surface forms.

This chapter focuses on the sound relations, beginning from the perspective of word-learning.

In the networked-morph model, the lexicon shows relations among morphs and phonological statements serve three functions: (i) to define wellformed morphs, (ii) to relate morphs in a morph set, and (iii) to select appropriate combinations of morphs in word formation. We discuss each in turn briefly, then provide deeper exemplification in §3.

17.2.2 Morph wellformedness conditions

Morph wellformedness conditions define what a phonologically wellformed morph is. They define the sets of possible segments from which morphs are composed, for instance whether there are [low, advanced] vowels, or only [low, retracted] vowels, etc. They also define wellformed sequences of sounds, for instance whether advanced vowels can precede retracted vowels and vice versa. They define wellformed structures, such as whether a syllable Onset can have more than one consonant, or whether advanced vowels are allowed in closed syllables, etc. We assume that wellformedness conditions are evaluated simultaneously for a given morph and so are largely surface-true.

Despite these being "largely surface-true," there are morphs that do not conform to all these conditions. We understand these exceptions in the context of Emergent Grammar, relying on similarity, frequency, and symbolic systems. Consider the origin of morph wellformedness conditions in a grammar. Schematically, a learner begins learning individual items, say fricatives: [s], [x], etc. These, and the morphs containing them, are sorted into bins according to various principles the learner posits, such as "fricative of type F," e.g. putting morphs with [s] in one bin, ones with [x] in another bin, etc. Attending to frequency, the learner may discover an asymmetry in these bins – for instance, the English learner will find many items in the [s] bin but very few in the [x] bin (we are assuming here that the learner has, on occasion, perceived [x]). Depending on the frequencies, the learner may decide the infrequent sound is rare but part of the language (e.g. [ʒ] in English), or may decide that the infrequent sound was a mistake and re-categorise members of that bin. A third alternative (which need not be co-temporaneous) is that the learner notes that all items in the [x] bin share properties – for instance they are all tokens of the word [bɑx] "Bach." In this case, the English learner would identify [bɑx] as a limited-frequency item that is counter to the general segmental conditions.

17.2.3 Morph relations

Simply having a set of wellformed morphs does not express all that a learner might identify using similarity and frequency. Based on semantic and syntactic similarity, a learner might group together pairs like [weɪt] and [weɪɾ] – both function as VERB; both mean "wait." As more items are encountered, there would be an increasing number of pairs which fit in the [t]/[ɾ] bin. At some point, the similarity of forms and the frequency of occurrence of the morph-final pattern would be generalised to encode how the morphs relate to each other, the second role of phonology in an emergent grammar.

The various bins and the number of items in each bin that a learner has identified give rise to numerous options when making a generalisation about the [t]/[ɾ] bin. For example, learners might (correctly) relate [t] → [ɾ]; they might also (incorrectly) relate [ɾ] → [t] or even [t] ↔ [ɾ]. In the latter two cases, as more items are encountered and sorted into bins, the learner would not only discover multiple items in the "exceptions-to-the-generalisation" bin, but also that the exceptions all share the property of pairing [d] and [ɾ] – leading to a readjustment of the grammar to the correct unidirectional generalisations. Similarly, on noting the [d]/[ɾ] cases, the learner who guessed right to begin with might modify the generalisation to a more general statement relating morphs with final coronal stops to morphs with final [ɾ].

Thus, morph relations state how two morphs are related to each other: there is no restructuring to create a single underlying representation from which the two morphs may be derived. Here we focus on phonological relations among morphs. Such relations can be unidirectional as with the English final coronal relating to a final [ɾ]. Such relations can also be bidirectional, for example the voicing alternation in a language like Imbabura Quechua where obstruent stops are voiceless except when post-nasal. In Imbabura Quechua, on encountering a voiceless stop, as in [wasi uku-pi] "inside the house," the learner can pair [-pi] with [-bi], positing a voiced stop, to use in [ñam-bi] "in the road." Similarly, on encountering [ñan-da] "road-ACC," the learner can posit a voiceless-initial suffix [-ta] to use post-vocalically as in [marja-ta] "Maria-ACC." (We ignore here the place alternation in [ñam-]/[ñan-] "road.")

We assume that morph relations express the minimal difference between two related forms and are expressed as implicational statements. Thus a relation [Low] → [Nonlow] has no effect on other properties such as [High] or [Back]. We also propose that each morph relation describes an observable relation between morphs: there is no ordering or ranking of morph relations. (This leads to some instances where the relations are either very complex or multiple relations function together as a set.) Further, these relations can only explicitly create a form that violates morph wellformedness conditions. Otherwise, satisfying the minimal change can only be achieved if it also satisfies morph wellformedness. In the absence of explicitly requiring violation of some sort of wellformedness, there could be no morph relation of a type producing "illformed" outputs.

17.2.4 Phonotactics and morphotactics

Since morphs are organised into sets, when multiple morphs combine to create a word, there can be several combinations to choose among. For instance, in creating the English past tense form *waited*, there are two verb morphs and three suffix morphs to consider. Combining { weɪt, weɪɾ } and { t, d, əd } results in six possible combinations to express the phonological form corresponding to WAIT-PAST:

weɪt – t   weɪt – d   weɪt – əd
weɪɾ – t   weɪɾ – d   weɪɾ – əd

How does the grammar determine which combination to use? Our answer is that the assessment is made by an interaction of phonological and morphological selection criteria as well as properties of the morphs themselves. For this English example, two phonotactics contribute to selecting [weɪɾəd]: (i) a preference for [ɾ] between stressed and unstressed vowels, and (ii) a preference against sequences of coronal stops.

In this chapter, we focus on the phonological properties of phonotactics and morphotactics; see Archangeli & Pulleyblank (2016) for discussion of morphological properties.

17.2.5 Summary

Under the networked model, morph sets contain related morphs. The role of phonological statements is threefold:

(i) morph wellformedness conditions determine whether a morph is wellformed and limit the effect of morph relations;
(ii) morph relations define how morphs relate to each other phonologically;
(iii) phonotactic conditions assess which combination of morphs is appropriate for a given context.

Before addressing how this model handles a range of phonological phenomena, we first exemplify these three phonological contributions, through an exploration of vowel patterns in Assamese (Mahanta 2012).

17.3 Case study: Assamese

The goal of this section is to demonstrate the function of each of the three types of phonological generalisations, using vowel behaviour in Assamese (Mahanta 2012).

17.3.1 Assamese morph wellformedness conditions

Assamese has an asymmetric set of eight vowels:

(3) Assamese vowels: { i, e, ɛ, ɑ, ɔ, o, ʊ, u }

Applying the tools of (1) to Assamese, the learner seeks to identify similarities among sounds, the more frequent sounds playing the most significant role in such identifications. Not only does the learner seek to identify sound chunks – such as the individual vowels in (3) – but the learner also seeks to identify similarities among these categories, such as the segmental conditions in (4a). Patterns in the distribution of the identified sounds are also identified, both positional, as in (4b), and sequential, as in (4c). (For focus, conditions here refer primarily to [ATR]/[RTR], or advanced and retracted tongue root. While we use familiar terms for describing these patterns, recall the discussion of (2), that these are not "universal features." Furthermore, the role of the segmental conditions is to define the segments of the language and to prevent the expectation of non-occurring segments, like *{ ɪ, ɑ̟, æ }.)

(4) Assamese segmental and sequential conditions (partial list)

a. Segmental
- i. [high, front] → [ATR]
  High front vowels are advanced, { i }.
- ii. [low] → [RTR]
  Low vowels are retracted, { ɑ }.
- iii. [low] → [back]
  Low vowels are back, { ɑ }.

b. Positional
- i. $*ATR_{nonhigh}\,C]\sigma$
  Nonhigh vowels in closed syllables are retracted.

c. Sequential
- i. $*RTR_{nonlow}$ˆATR
  Sequences of RTR_{nonlow}ˆATR do not occur.
  [i] is exempt.
- ii. *ATRˆRTR
  Within morphs, sequences of ATRˆRTR do not occur.
  [i] is exempt.

The segmental conditions of (4a) contribute to defining the possible vowels of Assamese in morph representations, whether learned directly or posited as a result of a morph relation.

Positional conditions of the type in (4b) serve two roles. First, the positional condition against advanced nonhigh vowels in closed syllables holds of morphs: [bɔnti] "lamp," [xɔrɔswoti] "Hindu goddess of learning," [kɛrketuwa] "squirrel," etc. One effect of this condition is that, while ATR harmony generally advances mid vowels ([kʰɔrɔs] "spend," [kʰorosi] "prodigal"), mid vowels in closed syllables remain retracted, as in [kɔrmɔ] "work," [kɔrmi], *[kormi], "active person." (We know of no cases in Assamese where affixation changes syllable structure so cannot test relevant predictions.) There are exceptions to this condition that are idiosyncratic: [nirdex] "command," *[nirdɛx]; [udbʰed] "exposure," *[udbʰɛd], etc. In Assamese, there are a small number of these unexpected forms: the learner must identify each one individually and learn that it is exempt from the positional condition on closed syllables.

The sequential conditions in (4c) both govern morph shape and selection among various morph combinations. In Assamese, while these are very strong tendencies, there are morphs that do not meet these conditions. Comparable to the items that have advanced mid vowels in closed syllables, there are morphs which have advanced-retracted sequences: [kolɑ] "black," *[kɔlɑ]; [perɑ] "sweet," *[pɛrɑ]; [bedɔnɑ] "pain," *[bɛdɔnɑ], etc. Again, each must be learned; there is a small class of such words. There are also items which have retracted vowels before advanced vowels as in [ɔxex] "limitless" (not *[ɔxɛx] nor *[oxex]) – which also violates the prohibition against advanced mid vowels in closed syllables.

To summarise, lexical items are expected to follow the morph wellformedness conditions in (4). When acquiring new items, the expectation is that these patterns will be followed, thereby facilitating acquisition. While there are exceptionless conditions, there can also be exceptions to very general patterns, as illustrated here with Assamese. In these cases, the representation of the exceptional morphs must override some general morph wellformedness condition. The expectation is that errors in learning these words will align with the conditions, and that language change will proceed in the direction of conforming to these conditions.

17.3.2 Conditions relating morphs

Learning Assamese will include learning items like those in (5), which show two morphs for each base.

(5) Assamese morphs

	observed items		glosses	morphs
a.	p^{h}ɛdɛlɑ	p^{h}edeli	"ugly M/F"	{ p^{h}ɛdɛl, p^{h}edel }
b.	gɛrɛlɑ	gereli	"fat M/F"	{ gɛrɛl, gerel }
c.	t^{h}ʊpʊkɑ	t^{h}upuki	"plump M/F"	{ t^{h}ʊpʊk, t^{h}upuk }
d.	pɑgɔl	pɑgoli	"mad M/F"	{ pɑgɔl, pɑgol }

Inspection reveals two ways these morphs could be related – morphosyntactically (one morph pairs with masculine and the other morph pairs with feminine) and phonologically (morphs with retracted vowels have counterparts with advanced vowels). As more Assamese forms are encountered, the relevance of the phonological relation is supported – pairs like [ʊtɔl] "boil" and [utoli] "boiling-INFINITIVE" show the same phonological relation but a different semantic relation.

The essential relation is that morphs with retracted vowels and otherwise identical morphs with advanced vowels correspond: { p^{h}ɛdɛl } ~ { p^{h}edel }, etc. The only exception seen is when the vowel is low, as in { pɑgɔl, pɑgol } "mad," where the requirement that low vowels be retracted (4a-ii) prevents [ɑ] from having an advanced counterpart. Does this mean the relation to be expressed is simply RTR ↔ ATR? We think not, for a variety of reasons.

First, in paired morphs with advanced vowels, all advanced vowels are followed by an advanced vowel, e.g. [k^{h}etori]. This suggests a unidirectional condition, RTR → ATR, supported by nonalternating forms like those in (6).

(6) Assamese ATR roots

a.	[b^{h}ut]	[b^{h}utɛ]	"ghost/ghost-ERGATIVE"	*[b^{h}ʊtɛ]
b.	[p^{h}ur]	[p^{h}urʊ]	"travel/travel-1ST.SG"	*[p^{h}ʊɛʊ]

Second, the alternating vowels are in a morph-final sequence: forms like [kɔpɑh] "cotton" and [zʊkɑr] "shake" have no corresponding advanced morph *[kopɑh], *[zukɑr]. Conversely, there are pairs where all vowels are either retracted or all are advanced: { gɔrɔm, gorom } "hot"; { bɛlɛg, beleg } "different." Thus, we propose RTR*] → ATR*], where "*" indicates a sequence of one or more vowels.

Finally, there is a special RTR/ATR relation with low vowels, as can be seen in (7): in these forms, the low vowel does have an advanced counterpart. Whether that counterpart is [e] or [o] depends on whether the low vowel follows a mid front vowel (7a) or not (7b–e). (See Mahanta 2012 for an alternative handling of these facts.)

(7) Assamese morphs with [ɑ]

	observed items		glosses	morphs
a.	dʰɛmɑli	dʰemelijɑ	"play/playful"	{ dʰɛmɑl, dʰemel }
b.	kɔpɑl	kopolijɑ	"destiny/destined"	{ kɔpɑl, kopol }
c.	gʊlɑp	gulopijɑ	"rose/pink"	{ gʊlɑp, gulop }
d.	sɑl	solijɑ	"roof/roofed"	{ sɑl, sol }
e.	pixɑs	pixosijɑ	"evil spirit/ill-natured"	{ pixɑs, pixos }

Thus, the general pattern RTR*] → ATR*] (8a) is not quite enough. If the final vowel is low, then the RTR/ATR relation is augmented by a low/nonlow relation (8b) or (8c), depending on the preceding vowel, if any.

(8) Assamese RTR/ATR relations
- a. RTR*] → ATR*]
- b. RTR*] → ATR*] & low] → nonlow]
- c. RTR*] → ATR*] & front, nonhighˆlow] → front, nonlow]

The RTR/ATR condition and the low/nonlow conditions cannot be collapsed together because the low/nonlow condition holds only over one vowel while the RTR/ATR condition holds over a sequence: [ɑlɑx] "luxury" and [ɑloxuwa] "pampered," *[oloxuwa].

These conditions identify relations between morphs. Let us consider how they function if a learner encounters the morphs { gɛrɛl } "fat" and { ɑlɑx } "luxury." In (9), the ✓ indicates that the requirements of the condition are met.

(9) Assamese morph generation

	morph	gɛrɛl	ɑlɑx	pixɑs	dʰɛmɑl
a. (8a)	RTR*] → ATR*]	✓	✓	✓	✓
b. (8b)	. . . & low] → nonlow]	–	✓	✓	–
c. (8c)	. . . & front, nonhighˆlow]	–	–	–	✓
	morph generation	gerel	ɑlox	pixos	dʰemel

The existence of the morph relations allows the learner to posit a second morph in each case. The morph sets that result from (9) are shown in (10); we use subscripted small caps to show the morphosyntactic and morphosemantic feature bundles that are associated with these phonological forms, schematised here with the glosses for the items.

(10) Morph sets resulting from morph generation in (9)
- a. $\{$ gɛrɛl, gerel $\}_{\text{FAT}}$
- b. $\{$ ɑlɑx, ɑlox $\}_{\text{LUXURY}}$
- c. $\{$ pixɑs, pixos $\}_{\text{EVIL SPIRIT}}$
- d. $\{$ dʰɛmɑl, dʰemel $\}_{\text{PLAY}}$

Note that in each case, the least change is made that can satisfy the conditions. For { gɛrɛl }, since both vowels satisfy the "final RTR sequence," but the low conditions are not satisfied, only (8a) is relevant: morph generation affects all vowels. A form like *[gɛrel] could not be posited for it does not satisfy the relation. Both { ɑlɑx } and { pixɑs } have the required final sequence of retracted vowels and both have a final low vowel. Morph generation results in paired morphs that satisfy both sets of requirements, differing only by the final vowel. With { d^hɛmɑl }, both vowels are part of a final chain of retracted vowels, satisfying the RTR/ATR part of the condition; the final vowel is both low and following a nonhigh front vowel, satisfying the low-&-front part of the condition: the generated morph { d^hemel } meets all requirements of (8c).

The principles of morph generation used here are (i) simultaneous morph generation for each set of morph relations; (ii) maximal interpretation of "*" – X* is the maximal uninterrupted sequence of Xs; and (iii) minimal change such that the conditions are satisfied – thus, morph relation (9b) has an impact only on the final vowel of { ɑlɑx }, not on both vowels.

Finally, we assume that the most frequent morph is the default morph (indicated by an underscore). The default morph is chosen when all else fails to make a selection. We assume the retracted morph is the default for Assamese morphs with both retracted and advanced morphs. We turn now to the criteria by which nondefault morphs are selected.

17.3.3 Conditions on morph selection

The impact of morph generation is morphs with multiple forms, e.g. { gɛrɛl, gerel }, { ɑlɑx, ɑlox }, etc. How does the grammar determine the phonological form that a particular morphosyntactic feature bundle will map to? We assume that the morphosyntactic features must manifest in the morphs selected. When the relevant morph sets contain multiple morphs and those morphs combine freely (see §2.4), how does the grammar determine which combination is right?

Answering this question is the focus of this section. We consider here two classes of conditions governing morph combination selection: general phonotactic conditions and morph-specific conditions, or morphotactics. (See Archangeli & Pulleyblank 2016 for discussion of selectional criteria with a strong morphological component, as well as further discussion of defaults.)

17.3.3.1 Phonotactic criteria

As seen in (4c), sequences of vowels that agree in tongue root position are preferred within morphs. This general property also holds of morph selection.

(11) Assamese phonotactic on tongue root
$*RTR_{nonlow}$ˆATR: avoid sequences of nonlow RTR followed by ATR

As a phonotactic, this condition governs the selection of morphs when polymorphemic forms are needed, eliminating sequences that do not agree for ATR. We show assessment of morph combinations in a table, very similar to an Optimality Theory tableau. The upper lefthand cell shows what is to be combined (here we show the morphs themselves; this could also show the morphosyntactic feature bundle, e.g. MAD-FEMININE for (12a)). The lefthand column presents the options under consideration, generated by all logical combinations of the morphs to be manifested. Selection is made by the conditions of the language, here (11).

In (12i), the sequential condition $*RTR_{nonlow}$ˆATR eliminates combination (12i-b) due to the disharmonic [ɔli] sequence. By contrast, the default form of the root is selected in both (12ii, iii) because there is no following ATR suffix.

(12) Assessments of Assamese morph combinations

i. {pagɔl, pagol}-{i}$_{\text{FEM}}$	*RTR$_{\text{NONLOW}}$^ATR	Default
☞ a. pagol-i		*
b. pagɔl-i	*!	

ii. {pagɔl, pagol}-{Ø}$_{\text{MASC}}$	*RTR$_{\text{NONLOW}}$^ATR	Default
c. pagol		*!
☞ d. pagɔl		

iii. {gɛrɛl, gerel}-{ɑ}$_{\text{MASC}}$	*RTR$_{\text{NONLOW}}$^ATR	Default
☞ a. gɛrɛl-ɑ		
b. gerel-ɑ		*!

More complex cases will involve multiple conditions; in some cases there is evidence for ranking of conditions. Note that there is nothing "exceptional" about these cases: regular phonological patterns of alternation are accounted for in this approach by productive generation of morph combinations with outputs assessed by general phonotactics.

17.3.3.2 Selection-by-morph, or morphotactic criteria

In addition to general phonotactic criteria, there can also be cases where specific morphs impose requirements on the phonological form.

We see an example of this in the behaviour of the low vowel in Assamese. In general, low [ɑ] is exempt from ATR harmony in Assamese (13). This is characterised by the sequential condition (11), which is restricted to nonlow vowels, i.e. *RTR$_{\text{nonlow}}$^ATR. Because of this restriction, the low vowels in (13) are unaffected by the suffix -i (as are the vowels that precede the low [ɑ]).

(13) Assamese low vowels do not harmonise

a. kɔpɑh "cotton" kɔpɑhi "made of cotton"

b. zʊkɑr "shake" zʊkɑri "shake-INFINITIVE"

However, as seen in (7), [ɑ] also alternates with an advanced mid vowel, with [e] when following a mid front vowel, else with [o]. This second pattern occurs before specific suffixes, in particular the adjectivisers [-ija] and [-uwa]. The generalisation, then, is that selection of the advanced morph is driven by the relevant (atypical) suffixes, encoded by stipulating that each requires a preceding ATR: { iyɑ$_{\text{ATR}}$__ }. Selection(-by-morph), a general requirement that selectional requirements outrank Default, gives rise to [kopoliyɑ] (14b) when affixed and the default [kɔpɑl] (14c) when no suffix follows.

(14) Assessments of Assamese morph combinations

i. {kɔpɑl, kopol}-iyɑ$_{\text{ATR}}$ __ }$_{\text{ADJ}}$	Selection	*RTR$_{\text{NONLOW}}$^ATR	Default
a. kɔpɑl-ijɑ$_{\text{ATR}}$__	*!		
☞ b. kopol-ijɑ$_{\text{ATR}}$__			*

ii. {kɔpɑl, kopol}	Selection	*RTR_{NONLOW}ˆATR	Default
☞ c. kɔpɑl			
d. kopol			*!

The sequential criterion *RTR_{nonlow}ˆATR has no impact because the root-final vowel in *[kɔpɑl-ijɑ] is low, so it is exempt from this condition.

17.3.4 Summary

This section has illustrated the three contributions from phonological generalisations: morph wellformedness, morph relations, and morph-selection criteria, both phonotactic and morpho-phonotactic. Morph wellformedness conditions define the phonological parameters that morphs follow for a given language. They typically hold over morph generation as well as over individual morphs. Morph relations define how morphs can be generated from each other. These relations are simply relations between morphs themselves; they do not take into consideration the contexts in which the various forms occur.

Morph wellformedness, morph relations, and selection-driving phonotactics can bear great similarities to each other. This raises the spectre of the "duplication problem." At its inception, generative phonology raised the duplication issue that the same regularities observed within morphs and across morph boundaries are characterised by two distinct statements within the grammar. The solution offered in generative phonology is to regard both effects as the consequence of a phonological rule, one that applies within both non-derived and derived lexical items. Within the model proposed here, phonotactics that hold of the sound system are expected to also govern morphs. However, some phonological relations have a restricted morphological domain; as such, not all phonological relations necessarily hold of all words.

Have we introduced a new duplication problem? Relations among morphs refer to phonological properties (for instance, coronal-final English morphs have a flap-final morph) and those same properties appear in phonological relations (flaps are preferred between stressed and unstressed vowels). Is duplication rearing its head? We contend not: the two types of relations are quite distinct from each other. Morph relations are drawn from observations about different surface forms that correspond to similar/same meanings but that have different phonologies: "coronal-final English morphs have a flap-final morph." Thus, morph relations express the phonological relations that may generalise over other morphs. Phonological relations define what is phonologically well-formed ("flaps are preferred between stressed and unstressed vowels"). The role of these statements is to identify well-formed morphs and morph combinations given the sound relations in each morph combination.

Like Optimality Theory, then, Emergence decomposes phonological relations, but the results are different than those found with OT. In Emergence, morph relations correspond roughly to the structural change of a rule-based model while the environment loosely corresponds to phonotactics in Emergence.

In this section we have dwelt more upon morph wellformedness and morph relations than on phonotactics and morphotactics; in the next section, we consider an Emergent account of a variety of phonological phenomena, shifting the focus to the criteria used in assessing among morph combinations.

17.4 Accounting for phonological patterns

We have presented conceptual motivation for exploring the Emergence hypothesis and a sketch of our current conception of Emergent Grammar. In particular, we have presented the morphological underpinnings we assume, with direct correspondence between the features specified in the morphosyntax and the morphs present in the morphophonology. We now turn to the very interesting question of how a variety of familiar phonological phenomena are handled under Emergence; in this discussion we see several ways in which selection of the appropriate morph combination is achieved.

17.4.1 Features and (un)natural classes: Tiv

Under the Emergence proposal, features of sounds emerge (Mielke 2008). Is this a reasonable position, given the high degree of similarity across languages with respect to features and natural classes? There are several types of responses to this question. First, because humans are physically quite similar to each other, it is not surprising that languages converge on very similar categories. Physical similarity – whether acoustic or articulatory – leads us to expect a high degree of cross-linguistic feature/natural class similarities while at the same time allowing for differences in details. Second, the similarity between classes cross-linguistically may be more of an assumption than a demonstration. Phonological descriptions often start with symbolic representations ("The vowels of this language are [i, e, a, o, u] . . .") where neither phonetic nor phonological evidence is actually presented for, say, [i, e] being similar to the categories symbolised in the same way in another language (Pulleyblank 2006). The third response is that "unnatural classes" occur in languages as well, a point demonstrated in Mielke (2008): in about a quarter of the large number of cases considered, groups of segments patterning together do not match up to distinctive feature proposals.

A case in point is the class of segments that occur in word-final position in Tiv, illustrated by a comparison between the Past and Recent Past tenses in Tiv (Arnott 1958; Pulleyblank 1988). Tiv Past has a final [a] or [e] vowel, while the Recent Past either has a final vowel that matches the preceding vowel (15a) or has no final vowel (15b). The latter occurs only with verbs with final { v, l, r, ɣ, m, n }.

(15) Tiv unnatural classes

	Past	Recent Past	Verb Gloss
a.	hèmbà	hèmbé	"exceed"
	ꜜcíngè	cíngí	"wind rope, etc. around thing"
	kùndè	kùndú	"mix things together"
b.	ꜜgéɣà	géɣ	"gulp"
	ꜜtírè	tír	"halt"
	ꜜbúmè	búm	"be foolish"

In attempting to characterise the class of final segments in (15b), Pulleyblank (1988, 315) describes the problem: it has been suggested:

> that the class is [+sonorant], but this is problematic for [v] and [gh] [= [ɣ] (da/dp)]. If they are analysed as sonorants, then a rule changing such sonorants into fricatives would be required. Perhaps a more plausible way to characterise this class is

to define it syllabically, that is, as those consonants that can appear in a (word-final) syllable rhyme.

In this type of case, similarity is derived by behaviour, not necessarily by acoustic or articulatory properties. The Emergent Hypothesis predicts that such classes can exist, but should be less common than classes where there are other similarities as well.

17.4.2 Autosegmental phenomena in Margi

Considering tonal phenomena from a more surface-oriented Emergent approach leads to very different analyses than in a more conventional autosegmental theory, sometimes in surprising ways. To illustrate, we consider one example from Margi; to explore the range of issues would be well beyond the scope of this survey.

As background, Margi (Hoffmann 1963) has two level tones, Low (v̀) and High (v́). Roots can be either L or H, as can suffixes:

(16) Margi tones

	Verb		H affix		L/H affix	
a.	tá	"cook"	tá-bá	"cook all"	tá-ná	"cook & put aside"
b.	ptsà	"roast"	ptsà-bá	"roast thoroughly"		
	ghàl	"grow old"			ghàl-nà	"wear out"
c.	fà	"take"	fá-bá	"take out"	fà-nà	"take away"

The examples in (16) illustrate three tonal root types and two suffix types. Roots may be (i) invariably H, e.g. tá "cook"; (ii) invariably L, e.g. ptsà "roast," ghàl "grow old"; (iii) L in some cases, H in others, e.g. fà ~ fá "take." Suffixes are comparable to the first and third root types: invariably H, e.g. -bá, and alternating between L and H, e.g. -nà ~ -ná. The morph sets for the items in (16) are shown in (17).

(17) Tonal morph types

	Tonal type	Roots	Affixes
a.	{ H }:	{ tá }	{ bá }
b.	{ L }:	{ ptsà }; { ghàl }	
c.	{ L ~ H }:	{ fà, fá }	{ nà, ná }

Two observations are of interest. First, if we consider the roots that have two morphs, one L, one H, we see that although there are two morphs there is no predictability: as the invariably { H } forms show, not all morphs with an H are paired with a morph that is L; and as the invariably { L } forms show, not all morphs with an L are paired with a morph that is H. Unlike the cases of multiple morphs seen in section 17.3.2, the Margi { L ~ H } cases simply have two lexically specified morphs. Second, and in contrast with the roots, we see that the tonal alternation in suffixes holds semi-predictably. There are no nonalternating { L } suffixes, only { H } suffixes and { L ~ H } suffixes. Hence for suffixes, there is morph redundancy of the type seen in section 17.3.2.

(18) Morph relation in Margi suffixes
[L] → [H] A Low-toned suffix has an H-toned morph.

This transparently represents an asymmetry in the implications of encountering an L or H suffix as a learner. If a learner encounters an H on a suffix, there is no way to tell if the suffix is invariably H or alternates between L and H. If a learner encounters an L on a suffix, then it is expected that the suffix will also occur with an H.

Surface forms in cases where the component morphs have single morphs are derived without any issues. Since there is a single morph, that is what occurs. Hence we observe words like *tá* "cook," *tá-bá* "cook all," *ptsà* "roast," and *ptsà-bá* "roast thoroughly."

Cases involving alternating roots like { fà, fá } "take" have options: for a given morpho-syntactic context, should the L-toned *fà* or the H-toned *fá* be used? Where there is no suffix the root occurs with an L tone, straightforwardly achieved by assuming that L morphs are the default (19a). If L morphs are the default, then when alternating roots are combined with alternating suffixes, the L-toned morph of each is expected, and this is exactly what occurs: *fà-nà*. Selection by default is illustrated in (20).

(19) a. { L } > { H }
b. { fà, fá }; { nà, ná }

(20) Margi assessment #1

i. { fà, fá } – { nà, ná }	Default
☞ a. fà-nà	
b. fà-ná	*!
c. fá-nà	*!
d. fá-ná	*!*

When alternating roots or affixes are combined with nonalternating morphs, the default gives way to the H-toned morph. To prevent having an LH or HL combination (*fà-bá, *tá-nà), we assume phonotactics prohibiting sequences of nonidentical tones:

(21) a. *HL
b. *LH

We see the effect of these phonotactics in sample cases in (22).

(22) Margi assessments (cont.)

i. { tá } – { nà, ná }	*HL	*LH	Default
a. tá-nà	*!		
☞ b. tá-ná			*

ii. { fà, fá } – { bá }	*HL	*LH	Default
a. fà-bá		*!	
☞ b. fá-bá			*

iii. { ptsà } – { bá }	*HL	*LH	Default
☞ a. ptsà-bá		*	

Sequences of nonidentical tones are ruled out – wherever possible, shown in (22i,ii). In a case like (22iii), a violation of *LH is inevitable because { ptsà } has a single L morph and { bá } has a single H morph. A form like the non-occurring *[ptsá-bá] is not even considered in (22iii): for this form to be a possibility, there would have to be an H-toned morph in the morph set containing { ptsà }, but there is no motivation to posit such a form since it does not occur. Finally, Default is not violated simply by the presence of an H morph (22i.a, ii.a, ii.b, iii.a); Default is only violated if a morph set has multiple morphs and a nondefault morph is present (22i.b, ii.b).

In these background cases, we see that regular Margi tonal alternations can be treated in a manner entirely comparable to segmental examples seen in section 17.3.3.

Consider now patterns involving classic elements of autosegmental representations (Goldsmith 1979).

(23) Margi tones

Verb		H affix		L/H affix
vǝ̌l	"fly, jump"	vǝ̀l-bá	"jump over, across"	vǝ̀l-ná "fly"

In (23), we have a word with a contour in one form (*vǝ̌l*), and a level tone in others (*vǝ̀l-bá*, *vǝ̀l-ná*). The classic autosegmental analysis views the alternating suffix *-nà*/*-ná* as underlyingly toneless and attributes the L.H pattern of *vǝ̀l-ná* to a distribution of two tones over two syllables where in vǝ̌l the two tones combine to form a rising contour. Under this view, the patterns of tonal identity seen with *-na* in (16) can be attributed to tonal spreading. In short, Margi has been taken to illustrate tonal melodies (L, H, LH, toneless), tonal spreading, contour formation, etc. How then do we account for such contour patterns in the emergent account presented here?

First, it is obvious that words like *vǝ̌l* appear in two surface forms: [vǝ̌l] and [vǝ̀l]. Second, contours are restricted to monosyllabic words. This limits both the location of contours in wellformed morphs and it limits where morphs with contours are permitted.

(24) Margi contour condition
If $[T_iT_j]_\sigma$ then $[\ \ ]_{Word}$ Contours appear only in monosyllabic words.

Third, there is a redundancy relation between these morphs. While the occurrence of an L morph does not at all mean that there is a corresponding rising tone morph (cf. *ptsà* "roast," *fà-nà* "take away," etc.), it does follow that if there is a rising tone morph then there is a corresponding L morph.

(25) Morph relation in Margi roots (first approximation)
$[\ \widehat{LH}\ ] \rightarrow [\ L\ ]$

Such a morph relation would correctly produce the two possible root forms, *vǝ̌l* and *vǝ̀l*, but would incorrectly predict that a suffix like *-na* should be L by default after *vǝ̀l*: **vǝ̀l-nà*, instead of *vǝ̀l-ná* (assuming the correct morph choice for the root). To obtain the correct result, we need to adopt the kind of *selection-by-morph* discussed in section 17.3.3.2. Instead of assuming that the morph related to a rising tone morph is simply L, we must assume that it is L and that it selects for an immediately following H (26).

(26) Morph relation in Margi roots (final)
$[\ \widehat{LH}\ ] \rightarrow [\ L\ ]\_\__{H}\ \}$

So the appropriate form for a morphosyntactic feature like FLY is: { *və̆l*, vəl __$_{\text{H}}$ }. In the absence of frequency data, we assume that the rising tone morph is the default in this case because it occurs in isolation; as seen in (27), however, assuming either form as default would work in this instance.

Forms such as *və̆l* and *vəl-ná* would therefore be derived as follows (where Select refers to the selectional requirement of the { vəl __$_{\text{H}}$ } morph and Contour refers to (24)).

(27) Margi assessments – contours

i. { və̆l, vəl__$_{\text{H}}$ }	Select	Contour	*HL	*LH	Default
☞ a. və̆l				*	
b. vəl __$_{\text{H}}$	*!				*

ii. { və̆l, vəl__$_{\text{H}}$ } – { nà, ná }	Select	Contour	*HL	*LH	Default
a. və̆l -nà		*!	*!	*	
b. və̆l -ná		*!		*	*
c. vəl__$_{\text{H}}$ -nà	*!				*
☞ d. vəl__$_{\text{H}}$ -ná				*	**

Interestingly, we see that the apparent autosegmental splitting of an LH sequence on one morph into a sequence of LH over two morphs need not be assumed in this approach. An LH on a single morph corresponds to an L morph selecting an H on a following morph. Whether this difference has benefits or disadvantages compared to a traditional autosegmental approach is unclear to us at this point. What is clear is that the basic machinery of the more surface-oriented emergent approach appears to account straightforwardly for the kinds of patterns we see in a "classic" tone system. In developing this account, the Margi example gives a flavour of how autosegmental phenomena such as melodies, spread, contour tones, etc. are accounted for under Emergence. These effects are the consequence of the interplay of phonotactic conditions governing tone sequences and morphotactic conditions governing morph types and the selection of tones on adjacent syllables, where each of the patterns is discernible from general tendencies in the surface forms of the language.

17.4.3 Abstractness in Tonkawa

According to generative and structuralist theories, morphemes have a single underlying form from which surface patterns are derived. Because these forms must be maximally information-bearing – encoding any phonological information that is not derivable by rule – there is a necessary class of cases where underlying forms must be posited that never occur in any form on the surface. Typically, this is because two properties α and β both occur in surface forms of some morpheme, but cannot cooccur in any surface form. In spite of their inability to cooccur in surface forms, both α and β must be posited underlyingly since they are not predictable. The need for such abstractness in generative theory and the lack of such abstractness in Emergent theory is discussed in Archangeli & Pulleyblank (2016). In this section, we present a specific example showing how the abstractness issue does not arise because the morphosyntactic features map directly to surface phonological forms, with morph relations expressed between these forms.

We examine Tonkawa (Hoijer 1933, 1949): as (28a) shows, biconsonantal verb roots have four different morphs: { pil, pile, pl, ple }. Since the form *pile* includes all information found in all surface morphs, it can be selected as the underlying form, /pile/, hence such biconsonantal forms are unproblematic and "concrete." The surface morphs depend largely on syllabification when various affixes are added. Triconsonantal verbs (28b) also have four morphs, e.g. { picn, picna, pcen, pcena }. Abstractness arises in this case because with triconsonantal roots no surface morph ever contains all three vowels. To encode the vocalic information present in all surface morphs, an abstract underlying representation like /picena/ appears necessary, even though it does not appear in any surface form.

(28) Tonkawa verb paradigm (roots in italics)

	Object	Present		Continuative	
a.	3SG	*pil*-oʔ	"he rolls it"	*pile*-n-oʔ	"he is rolling it"
	3PL	we-*pl*-oʔ	"he rolls them"	we-*ple*-n-oʔ	"he is rolling them"
	1ST	ke-*pl*-oʔ	"he rolls me"	ke-*ple*-n-oʔ	"he is rolling me"
b.	3SG	*picn*-oʔ	"he cuts it"	*picna*-n-oʔ	"he is cutting it"
	3PL	we-*pcen*-oʔ	"he cuts them"	we-*pcena*-n-oʔ	"he is cutting them"
	1ST	ke-*pcen*-oʔ	"he cuts me"	ke-*pcena*-n-oʔ	"he is cutting me"

Under Emergence, morphs correspond directly to surface forms. Identifying relations between morphs is based on both meaning similarity and sound similarity. There are also general phonotactics which hold in Tonkawa, governing roots and non-roots: there are no CCC (or longer) sequences; two-sided open syllables are not allowed, that is, * . . . V.CV.CV. . . . We also note that all roots are consonant-initial.

(29) Tonkawa phonotactics – directly relevant for root shape

a.	*CCC	There are no sequences of three consonants.
b.	* . . . V.CV.CV . . .	Sequences of simple open syllables are restricted.
c.	Initial-C	Roots begin with consonants.

In addition to such phonotactic conditions, there are regular shape relations among morphs. Note that these regularities are restricted to "shape" – to CV relations – they make no reference to vowel quality. Vowel quality is unpredictable and must be determined idiosyncratically for each morph. There are two basic relations among verb root morphs. First, if a vowel is observed in a root morph, there will be a corresponding morph without that vowel; second, if there is a consonant observed in a root morph, there will be a morph where that consonant is followed by some vowel.

(30) Tonkawa morph relations

a. *V $C_iV \rightarrow C_{i\emptyset}$

b. CV $C_iX \rightarrow C_iVX, X \in \{C, \#\}$

Consider the effect of these morph relations given various acquisition scenarios. With a biconsonantal root, if a form like . . . *pile* . . . was encountered, then the learner would posit { pile } – the specific form encountered – along with three additional forms resulting from the *V morph relation: { ple } (no V corresponding to the first vowel), { pil } (no V corresponding to the second vowel), and { pl } (no V corresponding to both vowels). This would give all four morphs: { pil, pile, pl, ple }. Imagine instead that a form like . . . *pil* . . . was encountered. Direct observation would establish { pil }; the *V relation would establish { pl }. What about the other two morphs?

While the CV relation might lead the learner to expect forms like [pilV] and [plV], the absence of information about the quality of the "V" would block the postulation of such morphs: morphs represent pronounceable forms but a vowel without specifications is not pronounceable. So from . . . *pil* . . ., only two morphs could be established. Encountering any form with the "missing" [e], i.e. [pile] or [ple], would allow the learner to complete the morph. The implication for acquisition – which seems necessarily correct – is that all members of a root morph set can only be learned if all appropriate morphs have been encountered, illustrating all the vowels.

Consider next the case of a triconsonantal root. First, as should be clear from the earlier discussion, there is no single form from which all morphs could be constructed. For example, if the learner encountered the root form . . . *picna* . . ., there would be no basis for postulating the medial vowel [e]. On the basis of . . . *picna* . . ., it would only be possible to postulate two morphs: { picna, picn }, the former based on observation, the latter on the basis of *V. It would not be possible to postulate a morph like *{ pcn } because that would violate *CCC (29a); it would not be possible to postulate a form like { picen } because there would be no basis on which to establish the vowel [e], nor *[pican] since the linear order of segments would not be respected and there is no evidence for changing linear order in Tonkawa.

To establish the full set of morphs, it is necessary for the learner to encounter some form with the medial [e] in addition to . . . *picna* . . . , for example, . . . *pcen* . . . or . . . *pcena* . . . . At some point this will happen: suppose that the learner encounters . . . *picna* . . . and . . . *pcen* . . . , both meaning " . . . cut . . . ." Knowledge about these morphs includes, among other things, the general precedence relations among sounds, for instance that in . . . *picna* . . . , *p* precedes {*i*,*c*,*n*,*a*} and in . . . *pcen* . . . , *p* precedes {*c*,*e*,*n*}, which collectively give *p*ˆ{*i*,*c*,*e*,*n*,*a*}. The learner also knows the root phonotactics in (29) and the morph relations in (30). How does the learner make use of this knowledge?

From observing . . . *picna* . . . , the learner would postulate a morph { picna }; similarly observing . . . *pcen* . . . leads to the morph { pcen } (31a,b). On the basis of { picna } and *V, the learner would postulate { picn } (31c), related by the absence of the final [a]. Alternative forms, also related by the absence of some vowel, are not possible morphs due to the root phonotactics, shown in (31d,e). On the basis of { pcen }, information about precedence relations in { picna, pcen }, and CV, { pcena, picen } would be postulated (31f,g). These differ from { pcen } by the added vowel (the CV relation); there are no changes in the precedence relations among segments in the morph; and *CCC and * . . . V.CV.CV . . . are satisfied. Note that **picena* (31h) would not be postulated as a morph since it would violate * . . . V.CV.CV . . . (29b) even though related to *picna* by CV.

(31) Tonkawa morph generation

{ picna, pcen }	perceived	*V	CV	*CCC	* . . . V.CV.CV . . .
☞ a. picna	✓				
☞ b. pcen	✓				
☞ c. picn		✓			
d. pcna		✓		*!	
e. pcn		✓		*!	
☞ f. pcena			✓		
☞ g. picen			✓		
h. picena			✓		*!

Overall, encountering . . . *picna* . . . and . . . *pcen* . . . in conjunction with the morph relations of *V and CV enables the learner to establish five morphs { picna, picn, pcen, pcena, picen }: four of these are seen in verbs (28b); additionally *picen* "steer, castrated one" is seen as a noun.

In this kind of case, there is no single morph from which all other morphs can be recovered. This situation is fairly common. Consider, for example, a prefix with both H- and L-toned morphs (neither is contained within the other) and the case of English flap/stop alternations (again, { hɪt } does not contain { hɪɾ } and vice versa). In short, the "abstractness" of generative phonology is concrete "business-as-usual" under Emergence.

To complete the Tonkawa analysis, we need a mechanism by which to choose which morph surfaces in any given case. This is straightforwardly achieved by syllable phonotactics. As is evident in (32), Tonkawa does not allow VV sequences, and there are no CC onsets or codas.

(32) Tonkawa phonotactics

a. *VV — Sequences of two vowels are not allowed/syllables have onsets.
b. *[$_\sigma$CC — Syllables do not begin with CC.
c. *CC]$_\sigma$ — Syllables do not end with CC.

In addition, we have not presented any discussion of how default morphs are identified in Tonkawa. Recall that the default appears when other conditions fail to make a selection. In some instances, the default may simply be determined by behaviour; in other cases it may be the most frequently occurring morph, or the most representative morph, e.g. the one sharing most features with all other morphs. For the Tonkawa case, being representative means that the default will be a longer morph as the longer morphs share more properties with all the shortest morphs. For an example like { picna, picn, pcena, pcen, picen }, being representative results in { pcena, picna, picen } as possible defaults. In the absence of frequency data, we assume that morph-initial consonant clusters are preferred (#CC . . . , . . . CC# > . . . VCCV . . .). Continuing our example of { picna, picn, pcena, pcen, picen }, this causes { pcena } to be preferred to { picna, picen }: { pcena } is identified as the default morph, which is supported by (33). (An alternative would be preferring no CC clusters at all, which would give the ranking { picen } > { pcena }, { picna }; conditions would need some modification. See Archangeli & Pulleyblank 2016 for discussion of multiple defaults.)

Syllable phonotactics coupled with Default identify the correct surface forms. In (33), syllable phonotactics eliminate three of the five options, and the prohibition on two-side open syllables eliminates the remaining combination (33e); [picnoʔ] is correctly selected.

(33) Tonkawa assessments for { <u>pcena</u>, picna, picn, pcen, picen } "cut"

CUT-3SG	*VV	*[$_\sigma$CC	*CC]$_\sigma$	* . . . V.CV.CV . . .	Default
a. pcena-oʔ	*!	*!			
b. picna-oʔ	*!				*
☞ c. picn-oʔ					*
d. pcen-oʔ		*!			*
e. picen-oʔ				*!	*

In (34), syllable phonotactics eliminate (34c) (either *[σCC is violated (**we.pic.nnoʔ*) or *CC]σ is violated (**we.picn.noʔ*)); * . . . V.CV.CV . . . again eliminates the combination with *picen*, (34e). Three combinations satisfy the syllabic restrictions (34a,b,d): default makes the choice, preferring (34a) over (34b,d).

(34) Tonkawa assessments for { pcena, picna, picn, pcen, picen } "cut," cont.

CUT-CONT-3PL	*VV	*[$_\sigma$CC	*CC]$_\sigma$	* . . . V.CV.CV . . .	Default
a. we-pcena-n-oʔ					
b. we-picna-n-oʔ					*!
c. we-picn-n-oʔ		<*!>	<*!>		*
d. we-pcen-n-oʔ					*!
e. we-picen-n-oʔ				*!	*

There is an alternative approach to analysing the Tonkawa patterns, found in Gouskova (2003, 2007), based on the interaction of syllable structure and feet, rather than the more linear/segmental approach taken here. Such an approach would require different morph set relations and different phonotactics than those given here, but the net result would be the same. Note that the very essence of the Emergent model allows multiple mutually compatible grammars for the "same" language; each learner constructs a model consistent with the available data.

To summarise, abstractness is not an issue under Emergence because morphs are only posited when they are consistent with morph patterns observed on the surface. For Tonkawa triconsonantal roots, there is no single surface form which contains all three vowels, so there can be no morph of that shape. Thus, there is no single morph from which all surface forms can be derived. This is not a problem since knowledge of observed forms in conjunction with rules governing morph relations allows the production of full morph sets.

17.4.4 Derivational opacity in Standard Yoruba

Derivational opacity refers to the class of phenomena in which surface forms are systematically inconsistent with some general pattern; "derivational" refers to the idea that in the course of a derivation, a later rule creates the environment for an earlier rule, but the earlier rule no longer applies. A case in point is the interaction between vowel harmony, consonant deletion, and vowel assimilation in Standard Yoruba. For a more complete discussion, see Archangeli & Pulleyblank (2015).

In Standard Yoruba, mid vowels agree in ATR/RTR with a following nonhigh vowel, illustrated in (35).

(35) Standard Yoruba vowel harmony: nonhigh vowels

V1↓/V2→	mid ATR	mid RTR	low RTR
mid ATR	ekpo "oil" olè "thief"	–	–
mid RTR	–	ɛsɛ̀ "foot" ɔbɛ̀ "soup"	ɛ̀kpà "groundnut" ɔjà "market"

Additionally, there is a pattern of [r] ~ Ø alternation when [r] is adjacent to a high vowel, with concomitant vowel assimilation when the second vowel is high (36a) and no assimilation otherwise (36b) (Akinlabi 1993).

(36) Yoruba [r] ~Ø alternations

a.	nonhigh-r-high	oríkì	oókì	"(praise) name"
		òrùka	òòka	"ring"
b.	high-r-nonhigh	ʃiré	ʃié	"play"
		èkùrɔ́	èkùɔ́	"palm kernel"

The interaction of the [r] ~Ø alternation, assimilation, and vowel harmony gives rise to the derivational opacity effect, shown in (37). Vowel harmony precedes [r]-deletion; [r]-deletion precedes V-assimilation. The net effect is a surface form that is disharmonic: [òrùka] alternates with [òòka], not *[ɔ̀ɔ̀ka], etc. The harmony pattern is derivationally opaque because there are derived surface forms which do not obey harmony.

(37) Derivational opacity in Standard Yoruba

Underlying representation	òrùka	
Vowel harmony	–	
[r]-deletion	òùka	
V-assimilation	òòka	*ɔ̀ɔ̀ka

Consider now the Emergent analysis. First, the harmony pattern can be expressed as a phonotactic prohibiting advanced nonhigh vowels before a retracted vowel.

(38) Harmony phonotactic
$*\mathrm{ATR}_{\mathrm{NONHI}}\ \mathrm{C}_{\emptyset}\ \mathrm{RTR}$

With respect to the [r] ~Ø alternations, the patterns illustrated in (36) must be expressed as two independent relations because assimilation occurs only when a high vowel follows [r] (39a), not when a high vowel precedes [r] (39b).

(39) Yoruba morph relations
a. $V_i\ r\ V_{j[\mathrm{HIGH}]} \rightarrow V_i\ V_i$
b. $V_{i[\mathrm{HIGH}]}\ r\ V_j \rightarrow V_i\ V_j$

On perceiving { òrùka }, the learner identifies the string $V_i r V_{j[\mathrm{HIGH}]}$, and morph generation results in a second form, { òrùka, òòka }. The relation between the two is fairly transparent, both in how it is expressed and in the shape of the two forms. The two forms of words with the requisite VrV sequences are straightforward morph relations.

Let us compare the Emergent analysis of Yoruba with the Emergent analysis necessary for Yoruba′, a language like Yoruba except that harmony holds of both morphs related by the [r] ~Ø relation – that is, Yoruba′ has no "derivational opacity" and so Yoruba′ has morphs like { òrùka, ɔ̀ɔ̀ka }. What we see is that the necessary morph relation is quite complex (40).

(40) Hypothetical condition to produce harmonic morphs
$V_i r V_{j[\mathrm{HIGH}]} C V_k \rightarrow V_l V_l C V_k$ where
a. V_l has the height and rounding of V_i and
b. if V_l,V_k are nonhigh, then V_l has the TR value of V_k

The morph created by the hypothetical (40) contains a segment that is a composite of two segments from the original morph, but that is not necessarily identical to any of the original segments. Compare (39) with (40). The relation in (39b) is simple, consisting entirely of a mapping of [r] onto ∅ in a given context. The relation in (39a) is slightly more complex, consisting of two relations: (i) mapping of [r] onto ∅, (ii) mapping of V_i onto V_j. By contrast, the relation in (40) involves three component relations: (i) mapping of [r] onto ∅, (ii) mapping of features of V_i onto V_j, (iii) mapping of features of V_k onto V_iV_j. Independent of issues of transparency and opacity, we see that the unattested pattern, the Yoruba′ pattern, is the pattern involving the highest level of complexity.

To summarise, under Emergence the simpler analysis, which combines [r]-deletion and V-assimilation in a single morph relation, is the attested analysis while the more complex one, combining [r]-deletion, V-assimilation, and harmony, is not attested. Simpler cases are significant because they are more likely to be identified, learned, and preserved while more complex cases are trickier to identify, and susceptible to mislearning and so to language change (Blevins 2004). The prediction is that the "derivational opacity" effect is preferred in the class of cases where the morph relation is more simply expressed than the corresponding "derivational transparency" relation.

17.4.5 Summary

While the Assamese example of section 17.3 illustrated how a grammar is put together, the analyses in this section give an indication of the types of analyses that are possible under Emergence, and demonstrate that the Emergence approach resolves a number of persistent challenges for linguistic theory.

(i) *Unnatural Classes* (Tiv): sound sets that function together in some pattern but that do not have a shared physical property. Under Emergence, this is possible because the function is sufficient to trigger learning the class; it is less likely because the evidence is less robust than a functional class that is also supported physically.

(ii) *Autosegmental Representations* (Margi): certain phonological patterns appear to require an autosegmental representation. The surface-oriented morph model advocated here provides a ready means of expressing those patterns without reference to autosegmental features.

(iii) *Abstractness* (Tonkawa): because there is no requirement of a single underlying representation, Emergence does not give rise to the abstractness problem. All morphs are concrete in that they appear unmodified at the surface.

(iv) *Derivational Opacity* (Yoruba): derivational opacity arises due to assumptions about how grammars work, in particular assuming a single underlying representation for each "morpheme." Derivational opacity is not an issue under the morph approach of Emergent Phonology.

17.5 Future directions

Given the newness of emergent approaches to phonological patterns – of which this chapter is one possible approach – virtually all areas of phonology and morphology demand research. In this final brief discussion, we consider similarities and differences with a theory of phonology, Optimality Theory, and one of morphology, Distributed Morphology. These areas define a large class of topics for further examination.

We have presented a model of an Emergent Phonology which makes use of ranked conditions, bringing to mind the ranked constraints of Optimality Theory (OT; Prince &

Smolensky 1993; McCarthy & Prince 1993a). There are significant differences. OT assumes a priori knowledge: the set of constraints is universal as is the general architecture of the model. But Emergence is not simply OT without universals.

First, the model of Emergence proposed here is built on a morph model. Consequently there is no single input form for each morpheme: the role of Gen is not to create an infinite candidate set but rather to create possible surface forms by taking the Cartesian product of the component morphs. Second, the conditions for assessing the possibilities are not universal, and so no broad typological claims are made based on the permutations of rankings. Typological effects are the result of what is common to human perception, cognition, and anatomy, not due to a common HLF. Third, Emergence requires no "Faithfulness" constraints because all morph combinations are fully faithful. By contrast, OT constraints are divided into two types, Faithfulness and Markedness. The role of Faithfulness is in counterpoint to Markedness: because Gen creates an infinite set of candidates, the grammar must include significant pressure to minimise deviation from the input otherwise all words would sound the same. Under Emergence the set of morph combinations is small because every morph is based on a surface form: there is no formal nor functional need for a class of Faithfulness conditions.

We have presented a model of an Emergent Phonology which is wholly integrated into the morphology of a grammar. This brings to mind theories such as Distributed Morphology (DM; Marantz 1997; Harley & Noyer 1999). The Emergent model is built on surface observations, with a minimal role for an innate HLF in morphophonology. DM is embedded within the HLF model, assuming a priori knowledge about morphosyntactic relations. This suggests that Emergence and DM are not compatible. However, much of the research in DM is focused on the morphosyntax of languages, and at this point seems quite compatible with many of our conclusions about morphophonology.

Under DM, there is no "lexicon" in terms of a part of the grammar where morphosyntactic operations take place independently of the syntax. However, there are "Vocabulary Items," which are inserted via a process called "Spell-Out"; Spell-Out takes place after the morphosyntax is complete. Our approach to Emergent Phonology is to determine the appropriate phonological form for a given morphosyntactic feature bundle; thus our starting point is consistent with a version of DM. One might interpret our lexicon, complete with these morph sets and conditions, as the source for Vocabulary Items.

A second question arises in how DM handles (morpho)phonological alternations. Our proposal is the parallel assessment of one or more concatenations of morphs, with assessment based on selection and phonological conditions. This contrasts with Embick (2010), which argues for a derivational approach to phonology within DM, claiming that a nonderivational approach is incapable of determining morph selection in cases where morph choice is not made for phonological reasons, that is, cases like the Assamese suffixes in section 17.3.3.2. As shown here, the correct assessment can be attained in a parallel, nonderivational fashion given appropriate assumptions about the nature of representations and conditions.

Exploration of how minimising assumptions about the HLF can interact with phonetic, phonological and morphological patterns forms a core area for further work. Moreover, testing claims computationally and experimentally are of significant importance.

17.6 Conclusion

We argue in this chapter that the Emergent approach leads to a morph model of grammar, where phonology plays three roles.

(41) Three roles for phonology
  a. Conditions determine the phonological wellformedness of morphs.
  b. Conditions characterise relations among morph sets.
  c. Phonotactic and morphotactic conditions select appropriate combinations of morphs in word formation.
    - Phonotactic conditions assess the phonological wellformedness of words.
    - Morphs encode idiosyncratic properties about their own phonological make-up.
    - Morphs may impose phonological requirements on adjacent morphs.

Phonological properties that are morphologically restricted are characterised by the form of a particular morph and by morphs with special selectional requirements. These properties of the Emergent model readily account for a variety of different types of problems, some of which have plagued rule-based and constraint-based models: (un)natural classes, one-to-many and many-to-one effects in tonal systems, abstractness of underlying representations, derivational opacity.

This chapter provides an overview of how to understand phonology as an Emergent system, building a phonological grammar without recourse to a priori phonological knowledge. This approach is learner-centred and bottom-up, making the minimal core assumptions that the human infant has memory and the ability to find similarities, attend to frequencies, and build a symbolic system. At this point, there is no appeal to an HLF for the types of phenomena under consideration.

17.7 Further reading

Deacon (1997) develops the hypothesis that the key element of human brain evolution is the ability to create symbolic representations, and that language is one consequence of this ability.

Boersma (1998) lays out the theory of Functional Phonology, formalising the tension between articulatory ease and the desire for clarity within an Optimality Theoretic model.

Lindblom (1999) argues against nativism and in favour of language as a “behavioral emergent,” focusing on the emergence of sound systems.

Bybee (2010) develops a usage-based approach to language, building on domain-general notions of categorisation, similarity, chunking, and association by contiguity.

MacWhinney & O’Grady (2015) present 27 chapters examining language from the emergentist point of view, covering basic language structure, change, typology, acquisition, and language and the brain.

References

Akinlabi, Akinbiyi. 1993. Underspecification and the phonology of Yoruba /r/. *Linguistic Inquiry* 24.139–160.

Archangeli, Diana, Adam Baker, & Jeff Mielke. 2011. Categorization and features: Evidence from American English /r/. In *Where Do Phonological Contrasts Come From?*, ed. by Rachid Ridouane & G.N. Clements, 173–196. Amsterdam: John Benjamins.

Archangeli, Diana & Douglas Pulleyblank. 2012. Emergent phonology: Evidence from English. In *Issues in English Linguistics*, ed. by Ik-Hwan Lee, Young-Se Kang, Kyoung-Ae Kim, Kee-Ho Kim, Il-Kon Kim, Seong-Ha Rhee, Jin-Hyuang Kim, Hyo-Young Kim, Ki-Jeang Lee, Hye-Kyung Kang, & Sung-Ho Ahn, 1–26. Seoul: Hankookmunhwasa.

Archangeli, Diana & Douglas Pulleyblank. 2015. Vowel harmony in emergent grammar: The case of Yoruba. In *Current Research in African Linguistics: Papers in Honor of Oladele Awobuluyi*, ed. by Ọlanikẹ Ọla Orie, Johnson Folorunṣọ Ilọri, & Lendzemo Constantine Yuka, 140–177. Newcastle upon Tyne: Cambridge Scholars Publishing.

Archangeli, Diana & Douglas Pulleyblank. 2016. Emergent morphology. In *Morphological Metatheory*, ed. by Heidi Harley & Daniel Siddiqi, 237–270. Amsterdam: John Benjamins.

Arnott, D.W. 1958. The classification of verbs in Tiv. *Bulletin of the School of Oriental and African Studies* 21.111–133.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boeckx, Cedric & Massimo Piattelli-Palmarini. 2005. Language as a natural object-linguistics as a natural science. *The Linguistic Review* 22.447–466.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Bybee, Joan L. 1999. Usage-based phonology. In *Functionalism and Formalism in Linguistics*, ed. by Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, & Kathleen Wheatley, 211–242. Amsterdam: John Benjamins.

Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Deacon, Terrence. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: W.W. Norton & Company.

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*, Linguistic Inquiry Monograph, volume 60. Cambridge, MA: MIT Press.

Gerken, LouAnn & Alex Bollt. 2008. Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development* 4.228–248.

Goldsmith, John. 1979. *Autosegmental Phonology*. New York: Garland.

Goldsmith, John. 2000. On information theory, entropy, and phonology in the 20th century. *Folia Linguistica* 34.85–100.

Goldstone, Robert. 1994. The role of similarity in categorization: Providing a groundwork. *Cognition* 52.125–157.

Gouskova, Maria. 2003. *Deriving Economy: Syncope in Optimality Theory*. PhD dissertation: University of Massachusetts at Amherst.

Gouskova, Maria. 2007. Dep: Beyond epenthesis. *Linguistic Inquiry* 38.759–770.

Hall, Kathleen Currie, 2009. *A Probabilistic Model of Phonological Relationships from Contrast to Allophony*. PhD dissertation: The Ohio State University.

Harley, Heidi, & Rolf Noyer. 1999. Distributed morphology. *Glot International* 4.3–9.

Harris, John. 1994. Monovalency and opacity: Chichewa height harmony. *UCL Working Papers in Linguistics* 6.509–547.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: The early stages. In *Constraints in phonological acquisition*, ed. by René Kager, Joe Pater, & Wim Zonneveld, 158–203. Cambridge, UK: Cambridge University Press.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.

Hoffmann, Carl. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.

Hoijer, Harry. 1933. Tonkawa: An Indian language of Texas. In *Handbook of American Indian Languages*, volume 3, 1–148. New York: Colombia University Press.

Hoijer, Harry. 1949. *An Analytical Dictionary of the Tonkawa Language*, University of California Publications in Linguistics 5. Berkeley and Los Angeles: University of California Press.

Hopper, Paul. 1998. Emergent grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, ed. by Michael Tomasello, 155–175. Mahwah, NJ: Lawrence Erlbaum Associates.

Lacerda, Francisco. 1998. An exemplar-based account of emergent phonetic categories. *Journal of the Acoustical Society of America* 103.2980–2981.

Lindblom, Björn. 1999. Emergent phonology. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, University of California, Berkeley.

MacWhinney, Brian & William O'Grady (eds.) 2015. *The Handbook of Language Emergence*. Chichester, UK: John Wiley & Sons, Inc.

Mahanta, Shakuntala. 2012. Locality in exceptions and derived environments in vowel harmony. *Natural Language and Linguistic Theory* 30.1109–1146.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4.201–225.

Maye, Jessica, Janet Werker, & LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82.101–111.

McCarthy, John & Alan Prince. 1993a. Generalized alignment. In *Yearbook of Morphology*, ed. by Geert Booij & Jaap van Marle, 79–153. Dordrecht: Kluwer.

McCarthy, John & Alan Prince. 1993b. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Technical Report #3. Rutgers University: Rutgers University Center for Cognitive Science.

Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

Mielke, Jeff, Diana Archangeli, & Adam Baker. 2016. Individual-level contact reduces phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language* 92.101–140.

Mielke, Jeff, Adam Baker, & Diana Archangeli. 2010. Variability and homogeneity in American English /ɹ / allophony and /s/ retraction. In *Variation, Detail, and Representation: LabPhon 10*, ed. by Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio, & Nathalie Vallée, 699–730. Berlin: Mouton de Gruyter.

Newell, Allen. 1990. *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Newport, Elissa & Richard Aslin. 2004. Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48.127–162.

Pierrehumbert, Janet. 2001. Exemplar dynamics, word frequency, lenition, and contrast. In *Frequency Effects and the Emergence of Linguistic Structure*, ed. by Joan L. Bybee & Paul Hopper, 137–157. Amsterdam: John Benjamins.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Pulleyblank, Douglas. 1988. Underspecification, the feature hierarchy, and Tiv vowels. *Phonology* 5.299–326.

Pulleyblank, Douglas. 2006. Minimizing UG: Constraints upon constraints. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero, & Michael Scanlon, 15–39. Somerville, MA: Cascadilla Proceedings Project. http://lingref.com/, document 1430.

Tesar, Bruce & Paul Smolensky. 1998. Learnability in optimality theory. *Linguistic Inquiry* 29.229–268.

18
Laboratory phonology[1]

Abigail C. Cohn, Cécile Fougeron, and Marie K. Huffman

18.1 Introduction

Over the past 30 years, *laboratory phonology* has developed with the core idea that how speech is structured, learned, and used is best investigated through experimental approaches and integrated methodologies. Laboratory phonology (LP) draws on theories and tools from various branches of the sciences to elucidate the linguistic, cognitive, and communicative nature of speech. Thus *laboratory* is used here in a broad sense, representing experimental approaches, and *phonology* is meant to include all aspects of the organizational structure of speech. We see the term *laboratory phonology* as roughly synonymous with *experimental phonology*. In this chapter, our aim is to introduce LP: its key questions, methodologies, and critical results.

The critical ingredients of the LP enterprise are the interplay between experimental work, broadly defined, and theorizing (theory development and testing), combined with methodological innovation, conducted in a collaborative, integrative, and multidisciplinary manner. The scholars working in this approach share an understanding of the central questions about the nature of speech and how they are best investigated, but don't necessarily share more specific theoretical views. Common caricatures of phonetics and phonology hold that phonetics doesn't use enough (linguistic) theory and phonology doesn't use enough (experimental) data; LP is an intellectual space where the strengths of each can complement the other. Ultimately, LP is agnostic on the relationship between phonology and phonetics and on the adoption of particular theoretic approaches to phonology (beyond rejecting theoretical assumptions that do not accord with empirical findings). By embracing, even demanding, creative and careful research using the broadest possible range of methodologies, work in LP illuminates the relations between the many different aspects of our knowledge and use of speech and language. At its outset, LP focused primarily on the use of phonetic methodologies to inform phonological questions (as discussed in section 18.2), but it has evolved into an expansive investigation of speech and signed language, integrating questions common to a variety of fields including phonology, phonetics, language acquisition, psycholinguistics, speech sciences, sociophonetics, and historical linguistics. In short, it has developed into the multidisciplinary study of speech as part of a linguistic, biological, and social system.

This broad perspective of LP is reflected in the mission statement of the Association for Laboratory Phonology:

> The purpose of the association is to promote the scientific study of all aspects of the phonetics and phonology of spoken and signed languages through scholarly exchange across disciplines. The association is an international body open to scholars world-wide, and is committed to the advancement and diffusion of knowledge on the phonetics and phonology of all human languages.
>
> *(http://labphon.org/, accessed June 16, 2014)*

This statement clearly articulates the collaborative, multidisciplinary nature that is part of the essence of LP, but it leaves undefined its purview and methods, and refers to *phonetics* and *phonology* without defining either. This suggests that there is a shared understanding of what phonetics and phonology are, which at some level is true. Referencing both phonetics and phonology also suggests that LP is an *approach* rather than a discipline. Yet, at another level, this leaves unanswered a complex set of questions about the nature of phonology and phonetics and their relationship.

For our purposes it will be useful to have working definitions of phonology and phonetics while noting that the boundary is not always clear. The terms *phonology* and *phonetics* are frequently used to refer to both the disciplines of investigation and the subjects investigated through those disciplines (following here Laver's definition of discipline vs. subject).[2] We focus here primarily on the latter, though we will have occasion to discuss the former.

Phonology is generally understood to be the study of what speaker/hearers implicitly know about the sound structures and sound patterns of their language, and thus is uncontroversially part of the linguistic grammar. Phonetics, on the other hand, is concerned with the production, acoustics, and perception of these structures and patterns, including dimensions under linguistic control. Phonetics and phonology interact: phonological structure is realized through phonetic means, which is generally referred to as phonetic implementation (although researchers disagree about whether this involves a procedural, directional relationship). It is also clear that phonetic considerations shape observed phonological patterns. That said, in this regard we refer the reader to Scobbie's (2007) elegant analogy addressing how phonetics and phonology can be considered distinct while at the same time lacking a sharp boundary:

> Rather, phonology and phonetics would have a transition zone, like a tidal shore ecosystem, which is defined by its dynamic transitions between seabed and land surface. Sea and land are (like cognitive and physical domains) categorically distinct, but the tides create a habitat in its own right. [. . .] Overlap does not imply loss of identity: the land and the sea are not the same and neither are phonetics and phonology.
>
> *(Scobbie 2007: 27)*

Researchers struggling with the lack of a simple definition of the boundary have often responded by assuming either a sharper, more discrete boundary or the absence of a boundary. This boundary dispute is one of the critical issues that led to the LP enterprise, the goal of which is to address the whole realm of knowledge about speech. Resolving the nature of the boundary between phonology and phonetics is not a prerequisite to effectively and usefully applying experimental approaches to the study of speech with the goal of elucidating questions of both phonology and linguistic phonetics. (For a recent review of this issue, see Cohn & Huffman 2014.)

The Association's mission statement is not specific about the methods of LP, beyond emphasizing that they are "scientific." This lack of specificity is surely intentional, in that methodological innovation and breadth are key ingredients of the enterprise. Yet, we think LP practitioners hold a shared vision of the critical importance of rigorous methods for data acquisition and analysis working in tandem with explicit hypothesis testing and model building. As noted above, *laboratory* is meant in a broad metaphorical sense, denoting experimental work conducted both in and outside the lab, drawing together an extensive and ever growing toolbox of diverse methodologies from linguistics and neighboring fields (see section 18.5). However, most fundamentally, it is *experimental* in the sense discussed by Ohala & Jaeger (1986, in the first explicit collective foray into experimental phonology), and also *empirical* in the sense defined by Chater et al. (2015):

> An experiment, then is simply the creation [. . .] of a situation in which crucial observations [. . .] may be made in such a way that they will be free from as many anticipated distorting influences as possible.
>
> *(Ohala & Jaeger 1986: 2)*

> It [the view of linguistics] is empiricist in the belief that the justification of a scientific theory must drive deep into the quantitative measure of real-world data, both experimental and observational, and it is empiricist in seeing continuity (rather than rupture or discontinuity) between the careful treatment of large-scale data and the desire to develop elegant high-level theories.
>
> *(Chater et al. 2015: 58)*

In section 18.2 we give some historical perspective on LP, and in section 18.3 we discuss the ways in which LP investigates the core questions of phonology. In section 18.4 we focus on LP's collaborative and multidisciplinary nature, reviewing some key insights that a LP approach makes possible. In section 18.5 we present an overview of the toolkit of experimental methodologies that embodies the LP approach, highlighting innovative methodologies. In section 18.6 we consider future directions.

18.2 Historical perspective

The term *laboratory phonology* was coined by Janet Pierrehumbert in the planning stages of the first LabPhon conference, which took place at The Ohio State University, in June 1987.[3] LabPhon 1 was co-organized by Mary Beckman and John Kingston, and among its central goals was bridging the distinct subfields and subcultures of phonology and phonetics. Since that time, 15 LabPhon conferences have been held roughly biennially in the US, Europe, New Zealand, and Asia and have brought together an increasingly large international community of scholars with diverse backgrounds. The successful development of the community working in a LP perspective led to the founding of the Association for Laboratory Phonology (http://labphon.org/) and the launching of *Laboratory Phonology*, the journal of the Association for Laboratory Phonology in 2010, both celebrated at LabPhon 12.

The history of LP can be roughly divided into two phases. The first phase, discussed in this section, encompasses LP's inception and early work revolving around the relationship between phonology and phonetics as understood by the two disciplines at the time (as exemplified by LabPhon 1–4). The second phase, discussed in section 18.3, developed since the mid-90s as the questions were increasingly defined in terms of speech as cognitive science, embedded in a broadly defined communicative system (as evidenced in work presented at LabPhon 5 and as discussed by Pierrehumbert et al. 2000).

The history of LP is in a sense the history of phonology in its relationship to (linguistic) phonetics. Over the course of the 20th century, phonology developed as a subject and discipline distinct from phonetics, with phonology increasingly understood to be about *language* and phonetics to be about *speech* (see Anderson 1985 and Goldsmith & Laks *to appear* for discussion of these developments). The need for the more integrated approaches called for by LP grew out of an increasingly sharp division between these two disciplines as they were practiced starting in the middle of the last century. On the phonetics side, this was a period of marked advances in the understanding of acoustics, made available by the development of the sound spectrograph during World War II, seen in the seminal work of Fant, Ladefoged, Stevens, and others (for a review of 20th century developments in phonetics see *Conference Proceedings: From Sound to Sense: 50+ Years of Discoveries in Speech Communication* 2004 and also Kohler 2000 and commentaries on that work by Ohala 2000; Laver 2000). With these advances came the development of phonetics as a modern science, including work on theories of speech articulation, acoustics, and perception, and the beginning of speech technology as a field. However, these investigations were not always related to or motivated by what phonologists considered to be the core questions about the nature of speech as a part of linguistic knowledge. On the phonological side, there were advances in formal approaches to phonology led by Halle & Chomsky (e.g., Halle 1959), in many ways encapsulated in Chomsky & Halle's (1968) *The Sound Pattern of English* (*SPE*) and, later, further developed in terms of sub-segmental and supra-segmental representations (see for example Goldsmith 1995). At times, the coherence of a formal system and theory-internal argumentation trumped an examination of data in its full complexity. The increasing separation of theoretical phonology and experimental phonetics led to a growing gap between these approaches to understanding the nature of speech, with discipline-internal definitions of the core questions, terminologies, and received methods of data collection.

This disciplinary chasm led many researchers interested in the nature of speech as part of a linguistic system to develop new experimental and collaborative approaches (see Ohala & Jaeger 1986; Beckman & Kingston 1990 for discussion of these points). The goal was to break down disciplinary divisions that impeded rather than enhanced the understanding of speech. In the introduction to the first LabPhon volume, Beckman & Kingston make a direct call for interdisciplinary collaboration, criticizing earlier work that assumed two distinct subcultures of phonology and phonetics *a priori*. They conclude that:

> the list of phenomena requiring such hybrid methods and models is much larger than hitherto supposed. We believe that the time has come to undo the assumed division of labor between phonologists and other speech scientists; we believe this division of labor creates a harmful illusion that we can compartmentalize phonological facts from phonetic facts.
>
> *(Beckman & Kingston 1990: 5)*

In thinking about the relationship between the subject of spoken language and the disciplines through which we investigate it, we are reminded of the story of the six blind wise men and the elephant. In brief, six blind wise men come across an elephant, and each touches a different part of the elephant: the trunk, a tusk, an ear, a leg, the belly, and the tail; and each has a very different interpretation: a thick branch, a hard pipe, a fan, a trunk, a wall, a rope. Only by putting together all the information of these narrower truths can they start to understand the elephant.[4] The goal of LP, in our view, is to allow each of the wise women and men to contribute their methodologies and insights to collaboratively come to understand the elephant; that is, the nature of speech as part of a cognitive, physical, and social system. This echoes Laver's (2000: 32) observation of the benefits of interdisciplinary research whereby we combine "two or more disciplines converging on a common problem to produce an outcome that is more than the simple sum of the parts."

This disciplinary ecumenicalism has remained a hallmark of LP, and the first LabPhon conference helped define and establish this effort. This charge to the developing LP community was of course framed within the theoretical assumptions of the time. For example, Beckman & Kingston (1990: 1) point out that papers in the LabPhon I volume collectively "address a more general issue, that of the relationship between the phonological component and the phonetic component." This topic, they say, encompasses three questions [emphasis added]:

> First, how, in the twin processes of producing and perceiving speech, do *the discrete symbolic or cognitive units of the phonological representation* of an utterance map into *the continuous psychoacoustic and motoric functions* of its phonetic representation?
>
> Second, how should the task of explaining speech patterns *be divided between the models of grammatical function* that are encoded in phonological representations and the models of physical or sensory function that are encoded in phonetic representations?
>
> And third, what sorts of research methods are most likely to provide good models *for the two components and for the mapping between them*?
>
> *(Beckman & Kingston 1990: 1)*

Notably, phonology usually meant "generative phonology" and the foundational assumptions that accompanied it. These included the assumption that phonology was language specific and phonetics was universal; that there was a sharp division between competence and performance; that the job of the generative grammarian was to describe and model the ideal speaker/hearer within a homogeneous speech community; and that lexical representations (which were also "underlying phonological representations") were sparse, encoding only contrastive information.

Phonological elements were defined as discrete and categorical, as opposed to continuous and gradient, with contrastive information encoded through bundles of distinctive features. Further, it was assumed that the grammar and its implementation were strictly modular, with a mapping between the two components, as noted in the italicized portions above. Several tenets of generative phonology were particularly influential in setting the common view of the relationship between phonology and phonetics. These views are laid out in Chomsky's (1965) *Aspects of the Theory of Syntax* and Chomsky & Halle's (1968) *The Sound Pattern of English* (see Cohn 2010 for discussion). Complementing these phonological views was an important research agenda within phonetics searching for "invariance" in the acoustic signal (e.g., Blumstein & Stevens 1981). On a simple modular view, perception was presumed to involve detecting the discrete phonological features in the speech input, so it was hypothesized that invariant cues (albeit possibly somewhat abstract ones) must be present and discoverable in the physical signal (see Lindblom 1990 for discussion of this issue).

There was also a widely shared view that perception was categorical (e.g., Repp 1984; see Gow & McMurray 2004 for discussion). Together these lines of research in phonetics were in harmony with, and subtly reinforced, the phonological view that contrast was represented categorically and locally.

Over the past 30 years, each of these positions has been rethought in part because of new experimental approaches enriching our empirical foundation, in turn leading to a more nuanced understanding of these points. However, each of these views, either explicitly or implicitly, still exerts some influence on the way many of us investigate, analyze, and interpret findings about speech, a point we return to in section 18.4.

Early work in LP incorporated developments in linguistic phonetics, termed "generative phonetics," building on Pierrehumbert's (1980) seminal dissertation in which she modeled phonological and phonetic aspects of intonational patterns in English with an extension of autosegmental representations. This, along with work on segmental patterns (such as duration, nasalization, and consonant-to-vowel and vowel-to-vowel coarticulation) by Keating and others (e.g., Keating 1985) evidencing the language-specific nature of phonetic realizations, resulted in a widely shared understanding that many aspects of phonetics were necessarily under speaker/hearer control. This line of work led to a rejection of the *SPE* view that phonology was part of the speakers' knowledge and phonetics was the automatic "universal" (extragrammatical) manifestation of that knowledge. An understanding of the commonalities and differences across the languages of the world continues to be a central concern of both phonology and phonetics, testing more nuanced theories of what might be "universal" and why. For recent discussion of this point, see Chater et al. (2015). LP has also contributed significantly to a deeper and richer understanding of the nature of variation, including fine details in the signal as it encodes not only linguistic meaning, but also other dimensions of communicative and social meaning. This body of work has substantially challenged the assumptions of sparse lexical representations and sharp division of competence and performance.

The theoretical significance of fine phonetic details for phonological and lexical representations grew out of methodological developments resulting from the increased awareness of the limitations and insufficiencies of impressionistic "narrow" transcription at the level of the "phone."[5] On the methodological side, such transcription does not include fine enough detail to serve as the "raw" data (or as the only source of raw data) for phonological analysis (see Ladd 2011). While Ohala & Jaeger (1986: 2) frame in broad terms the importance of the experimental method as "based on the recognition that our knowledge of the world is subject to many distortions," this call applies very directly to the widely used methodology in phonology of impressionistic transcription based on careful listening and introspective judgments. Depending on written, qualitative realizations of speech as our principle source of data is limiting because of the intrinsic biases that go along with them. We highlight some critical aspects of these biases. First, speech is continuous, and transcription, no matter how careful and narrow, is discrete. A fundamental property of perception of speech is that it is perceived in part categorically, so we as listeners are not consciously aware of the gradience of speech or how the language background and experience of the listener necessarily warps their perception. Since researchers vary in their sensitivity to the fine details of speech, and in the degree to which they are aware of these distortions, the interpretation of impressionistic data included in others' work is particularly difficult to interpret reliably. This does not mean there is no room in phonology for impressionistic transcriptions and introspective judgments; they serve as an excellent analytic starting point, but not an end point. They need to be augmented by experimental data, whether through signal analysis, behavioral judgments (as simple as testing of nonce forms – "wug" testing), or through more elaborate methods.

Scobbie (2007: 18) highlights the complexity that the fine-grained details of production and perception bring to the task of understanding what is phonologically relevant:

> Phonetically detailed studies of multiple speakers reveal the extent of language specific control of phonetic targets (often resulting in subtle interspeaker variation) in phenomena that are firmly within the phonological canon. Such work shows the extent to which subtle, gradient, and variable (i.e. phonetic) patterns exist alongside the gross and categorical (i.e. phonological) ones previously easily detected via native speaker intuition and impressionistic transcription of individuals or small homogeneous groups of speakers.
> *(Scobbie 2007: 18)*

Investigating and modeling these finer grained elements is at the heart of the LP research agenda. In our opinion, this increasing interest in phonetic details is not a simple switch of methodological lens, but is rather a change in the way we *do* phonology: LP has contributed to our understanding that there is not an *a priori* level of granularity that defines the elements of speech that are relevant for phonology. Thus, as experimental and quantitative approaches have enabled researchers to investigate more systematically the fine details of speech, this in turn has led to a rethinking of a number of key operational principles.

As LP became more interdisciplinary and drew in more participation from researchers in neighboring fields, the questions of LP were redefined more generally as questions about cognitive and biological systems, embedded in a broadly defined communicative system, as discussed in the next section.

18.3 Critical issues and questions

In this section we frame the questions of phonology as seen through the LP approach and consider how LP has changed the way we do phonology.

Ohala (2007: 3) states that:

> Broadly speaking, a scientific discipline can be characterized by:
>
> - the questions it asks;
> - the answers given to the questions, that is, hypotheses or theories;
> - the methods used to marshal evidence in support of the theories.

He argues that the broad questions asked in phonology (including, for Ohala, phonetics) have remained "remarkably constant over time." We find Ohala's list comprehensive and germane, and we quote here five of his eight questions, which define common concerns of phonologists interested in the LP approach:

1. How is language and its parts, including words and morphemes, represented in the mind of the speaker; how is this representation accessed and used? How can we account for the variation in the phonetic shape of these elements as a function of context and speaking style?
2. How, physically and physiologically, does speech work – the phonetic mechanisms of speech production and perception, including the structure and units it is built on?

3 How and why does pronunciation change over time, thus giving rise to different dialects and language and different forms of the same word or morpheme in different contexts? How can we account for common patterns in diverse languages, such as segment inventories and phonotactics? [. . .]
6 How is speech acquired as a first language and as a subsequent language?
7 How is sound associated with meaning?

(Ohala 2007: 3–4)

Solé et al. (2007: vi) summarize factors responsible for the increased use of experimental methods in phonology. The list is an elegant encapsulation of key issues that place experimental methods at the center of current research in phonology.

> First, phonology is addressing increasingly diverse questions about the structure of grammars and the representation of sound patterns in the mind and brain, about the relation between phonetic and phonological constraints, about categorization of sensory data, sound change, socially and geographically indexed variation, and so on; [. . .]
>
> Second, technologies relevant to phonological inquiry continue to evolve, as does the availability of large scale linguistic corpora; new technologies and databases open up new opportunities, new questions, and new grounds on which to test hypotheses.
>
> Third, there is growing recognition that phonological inquiry should be embedded within a framework informed by the biological, social, and cognitive sciences; application of standardized experimental techniques from these disciplines allows us to account for phonological structure in ways that are both consistent with established knowledge in these fields and (arguably) better able to provide a unified account of language and speech.
>
> Fourth, a clear demonstration that we understand phonetic and phonological principles is the ability to model relevant behaviors and patterns; consequently, the use of articulatory synthesis, stochastic methods, learning algorithms, pattern recognition techniques, and neural networks are of increasing importance to phonological inquiry.
>
> *(Solé et al. 2007: vi)*

With its strong commitment to experimental approaches, the LP community shares all of these interests and goals. These are (paralleling the four items in the list): broadening the purview of phonology, innovation in methodology, understanding phonology as part of cognitive science, and privileging hypothesis testing and data modeling. While there is debate about the boundaries of phonological inquiry, LP highlights several important points. First, there is no simple delineation between the core questions about the nature of representation and patterning of speech sounds (which we expect contributors to this volume would agree is "phonology"), and the embedding of a phonological system within a social and communicative system. LP comes down on the side of this broader purview. Second, progress on these questions has been achieved through rethinking "how" we investigate them, drawing on a wide range of quantitative and experimental methodologies. The result of this commitment to broad and rigorous methods has in turn led to hypothesis testing and modeling of data in innovative ways that both offer new insights and expand the kinds of questions we can ask about sound patterns and speech.

Through collaborations that evolved naturally with the increased richness of methodologies came a widening of the lens through which speech is investigated and understood. As noted above, starting explicitly with the themes of LabPhon 5 (which took place in 1996), the enterprise of LP was framed in terms of cognitive and biological systems (see Pierrehumbert et al. 2000 for development of this view). The expanding community of scholars led to greater interchange and collaboration between phonologists, phoneticians, psycholinguists, speech scientists, socio-phoneticians, and acquisitionists, resulting in greater attention to how language and speech are situated in a broader system, and recognizing the *linguistic*, *cognitive*, *communicative*, *social*, and *motoric/physiological* aspects of speech. These developments are nicely documented by Pierrehumbert & Clopper (2010: 114–117) in a network analysis of citations of work in LP. They show first the rich array of fields in which laboratory phonologists publish. They then show network linkages across work by these researchers, with bidirectional linkages across the core fields and unidirectional linkages with more peripheral fields "suggesting the transport of knowledge across disciplines." Pierrehumbert & Clopper (2010: 116) conclude: "Thus, laboratory phonology is not a subfield of linguistics or cognitive science, but a federation of scholars with similar research interests and goals." They see the primary contributions tying the federation together as the "relationship between the physical reality of speech and the cognitive representation of language" (116), "autosegmental-metrical phonological theory" where "hierarchical data structures from different time scales are analyzed in a single theoretical framework" (119), and increasingly "the relationship between levels of representation and the social function and social context of language" (117). (See also Cohn 2010; Croot 2010 for discussion of ways LP has developed.)

LP's richer empirical foundation, together with the ecumenical approach to disciplinary boundaries, leads to a more nuanced understanding of speech. The systematic examination of the fine details of production, acoustics, perception, storage, acquisition, and learning has not only changed the way we do phonology, it has also fundamentally shifted our understanding of the substance of phonology. The scientific enterprise requires us to move beyond theoretical assumptions that do not accord with this empirically grounded understanding of speech and language. LP has produced much evidence that various compartmentalizations, so common to most theories of phonology, turn out to be oversimplifications. Crucially, LP doesn't reject them as wrong, but rather recognizes them as overly simplistic, and strives to build on what these approaches have taught us. As exemplified in section 18.4, we conclude that by privileging new integrated and multidisciplinary approaches, LP has shed new light on the questions of phonology, offering new answers to old questions while also posing new questions.

18.4 Current contributions and research in laboratory phonology

In this section we exemplify, through a selection of results, how the methodological premises of LP have profoundly impacted how we do phonology. These include attention to fine phonetic details, prosodic representations and realization in relationship to segmental properties of speech, multifaceted dimensions of meaning including socio-indexical meaning, and multimodal information used in speech processing. For a more comprehensive sense of the breadth and depth of the work done under the LP umbrella, we refer the readers to the LabPhon volumes I–X, the journal *Laboratory Phonology*, and *The Oxford Handbook of Laboratory Phonology* (Cohn et al. 2012).

18.4.1 Language-specific phonetic detail

In the early stages of LP research (see section 18.2), much research focused on understanding the division between phonology and phonetics, interpreted as an opposition between categorical and discrete, and continuous and gradient, phenomena. The accumulating data on the finer phonetic differences between sounds that might be considered phonologically "the same" in different languages was particularly challenging for this presumed division. The documentation of fine phonetic details of speech informs our understanding of how phonemic contrasts are realized in the signal. This large body of literature in fact caused a major shift in perspective, revealing that, contra longstanding assumptions, phonology is at most quasi-discrete, quasi-categorical, and quasi-local.

Many early LP studies took this problem head on, and investigation of a wide range of phonological processes originally described as categorical established that these processes are more continuous and gradient than previously assumed (see Ernestus 2012 for a review). For example, Nolan (1992) showed that the assimilation of a final /d/ to the place of articulation of a following word-initial velar in English does not result in a categorical shift in identity of /d/ to /g/. Indeed, acoustic and linguopalatal contact properties of word pairs like *lead* /lɛd/ and *leg* demonstrated that in the assimilated context *lead* presented a continuum of realizations, ranging from forms with full alveolar contact to forms with no alveolar contact, but always distinct from *leg*. Similar evidence against complete categorical neutralization of contrast has been found in an array of different languages (e.g., palatalization in Russian, Barry 1992, and in English, Zsiga 1995; assimilation of place of articulation in English, Gow 2001; Ellis & Hardcastle 2002; final devoicing in Dutch, Warner et al. 2004; schwa elision in French, Bürki et al. 2011). These studies have thus refined the widely held assumption that contrastive information is realized in discrete ways.

Speaker/hearer awareness of these finer details and sensitivity to (token) frequency in both perception and production have led to widespread rejection of the view of lexical representations containing only contrastive information (see, e.g., Coleman 2003). Effects of token frequency and various priming effects in lexical access have been widely observed, but less well understood is the degree of such effects in production (see Jurafsky 2003 for a review). Observations of sensitivity to fine details, and their incorporation at least in short-term memory, have contributed to the development of exemplar models of speech perception and production (e.g., Johnson 1997; Pierrehumbert 2001; see also Silverman 2011 for a review of usage-based approaches in phonology).

Similarly, experimental investigation of the speech signal has led to a more refined view of the way information associated with phonological contrast is distributed in time in the speech signal, calling into question traditional assumptions of locality. While contextual effects (coarticulation) have been well documented in the phonetic literature, they have often been attributed to effects of the speech apparatus (e.g., co-production) and have been left outside of the scope of phonology. (However, see Hoole et al. 2012 for an elegant accounting of how tightly grammatical and mechanical effects can be interleaved.) Long-distance coarticulatory effects raise questions about the assumed locality of scope of phonological representations and highlight the relatively weak account of the temporal domain given in classical phonological models (though see work in Articulatory Phonology for an approach integrating temporal relationships into phonological representations; e.g., Browman & Goldstein 1990; Gafos & Goldstein 2012). Some striking examples of how contrastive information is distributed broadly in the speech signal are discussed by West (1999); Coleman (2003); Hawkins & Nguyen (2004); and Hawkins (2012). For instance, the contrast between a pair of words like /lɛd/ and /lɛt/, which is abstractly localized on either the final voiced or voiceless consonant, is in fact realized with acoustic information distributed in time such that the word-initial /l/ shows different durational and spectral properties in the two words.

18.4.2 Prosodic representations and realization

An even more extreme example of non-locality in phonology is illustrated in studies concerned with the mapping between abstract phonological tonal units and the continuous and gradient speech signal. From its outset, the development of LP was very strongly driven by the work done on the prosodic structure of speech (considered as one of the three major "commodities" of LP by Pierrehumbert & Clopper 2010). One of the core questions in Intonational Phonology (Pierrehumbert 1980; Pierrehumbert & Beckman 1988; Ladd 2008; Grice 2006; Arvaniti 2012) is the mapping of phonologically relevant tonal targets with segmental strings. As explained by Arvaniti (2012), in early autosegmental work (e.g., Goldsmith 1976) this relationship was assumed to be straightforward: elements that associate in phonology co-occur in time. Yet, detailed acoustics analysis has revealed that the phonetic realization of phonologically specified tonal targets does not always co-occur or synchronize with the associated segmental tone-bearing unit. Based on cross-linguistic and cross-dialectal observations, Ladd and colleagues (e.g., Ladd 2006; Ladd et al. 1999, 2000) have developed the idea of stable anchoring of tonal events with specific segmental or syllabic landmarks. Although the existence of "strict" segmental anchoring and the selection of the putative anchors for tonal alignment have been much debated in the literature, the notion of anchoring has not been abandoned, and recent studies have argued that the anchors should be found in the articulatory signal rather than the acoustic signal (see D'Imperio 2012).

Looking at speech in a broader context in LP is also demonstrated by work showing the prosodic coherence and possible alignment of speech and non-speech gesture. Loehr's (2012) study of free conversation found that for a broad range of gestures, gestural apexes aligned with pitch accents, and gestural phrases were often aligned with intermediate phrases. Esteve-Gibert & Prieto (2013) found that pointing gestures and intonation peaks show parallel behavior and are both constrained by prosodic structure. Further, Esteve-Gibert et al. (2014) showed that the apex of head gestures aligns with accented syllables. For additional discussion, see Wagner et al. (2014), and papers in the special issue of *Speech Communication* discussed there. Besides phrase-level phonological processes, finer grained phonetic variation has also been shown to inform and be informed by prosodic structure. Within the last 30 years, an extensive body of acoustic and/or articulatory evidence has shown that the phonetic nature of words, and their constituent segments, depends on their position relative to prosodic constituent boundaries and prosodic prominence. For instance, the glottalization of word-initial vowels in American English has been found to be more frequent at the Onset of intonational phrases (Pierrehumbert & Talkin 1992) and in pitch accented syllables (Dilley et al. 1996). The presence and strength of a prosodic boundary or prominence also affects other supralaryngeal articulatory details. Fougeron & Keating (1997) investigated variations in the amount of contact between the tongue and the palate for consonants and vowels at different levels of prosodic organization. This study showed that prosodic domain-initial /n/s were articulated with more contact than internal ones, and that the degree of contact was greater for consonants at the beginning of prosodic constituents higher in the hierarchy (e.g., intonational phrase vs. intermediate phrase). Cho (2011) provides an excellent review of experimental evidence of prosody-related segmental variations that are either boundary related or prominence related. These effects have been observed in many unrelated languages (English, French, Korean, German, Tamil, Japanese, and others), applying to both spatial and temporal parameters of the laryngeal, lingual, velic, jaw, and lip subsystems. (See also contributions in the recent issue of the *Journal of Phonetics Dynamics of Articulation and Prosodic Structure* 2014.)

18.4.3 Broader (re)definition of meaning

Another area in which the broader perspective of LP has been beneficial relates to a conceptual redefinition of linguistic meaning. From the diversity of phonetic details and factors influencing pronunciation that have been empirically documented, it is clear that lexical contrastive meaning is only one part of the information conveyed by the speech signal and that other types of meaning should be attended to in phonological investigation. As Local (2003: 322) pointed out, "meaning is much more than lexical meaning." In the early formulations of generative phonology, it was assumed that the domain of the grammar included only lexically contrastive sound properties, plus those contextually defined regular variations that could be codified as allophones. This early view privileged lexical contrast, and citation forms or at least relatively careful speech. While psycholinguistic research still supports the importance of this kind of knowledge (e.g., Sumner & Samuel 2005), it is at this juncture impossible to make a convincing case that linguistics is concerned with only this type of linguistic knowledge. The many language-specific aspects of subphonemic variation, whether contextually determined or socially indexed, have made clear that many more kinds of knowledge are learned and are critically involved in all activities in which humans use and acquire language (e.g., Foulkes & Docherty 2006). A central goal of LP, therefore, is to understand how to embed a theory of lexical/contrastive meaning into a broader theory of meaning, which places speech at the core of a wider system of communication.

Key to this broader view of meaning is a richer understanding of the nature of variation, including not only message-related variation (traditionally studied in phonology and linguistic phonetics) and speech-apparatus related variation (traditionally studied in experimental phonetics and psycholinguistics), but also speaker-, pragmatic-, interactional-, and discourse-related factors. Among those, a large number of studies have shown that word pronunciation is affected by lexical properties such as frequency of use in a language, and relative frequency compared to lexical neighbors or neighborhood density (e.g., Wright 2004; Watson & Munson 2007, but see Gahl 2015). For example, non-frequent words or words with numerous competitor neighbors have been found to be more hyperarticulated than others. The word's function, its role in discourse, or its predictability within context, have also been found to affect pronunciation (e.g., Bell et al. 2003; see overview in Ernestus 2012). Kelly & Local (1989), for instance, have shown that while assimilation of place in English is frequent for /n/ and less so for /m/, in the grammatical chunk *I'm*, assimilation regularly occurs in everyday speech, as in *I[ŋ] going*. Similarly, Local (2003) reports that word-final stop release is used to signal turn-taking in conversation. If phonetic details of words are found to vary depending on properties linked to their use, we have to understand how parts of phonology are influenced by, and interact with, the lexicon and communicative or pragmatic functions.

18.4.4 Social-indexical variation

Another type of variation that has more recently been an important focus in LP is variation in social-indexical information, which, interpreted most broadly, involves all variation correlated with non-linguistic factors (Abercrombie 1967; Foulkes 2010). Non-linguistic factors contributing to variation range from gender, age, and regional background to socio-economic status, group affiliations, and emotion. Social-indexical variation poses some of the same challenges to phonological theory that language-specific phonetic detail in phoneme contrasts has presented. Like fine linguistic phonetic (subphonemic) detail, social-indexical linguistic variation, which is associated with specific social groupings and individual identities, clearly is learned, is fundamentally defined in terms of linguistic units, and does not follow automatically from any physical or social principles (Foulkes 2010). A number of researchers in LP have recently discussed the essential link between traditional/lexical linguistic and sociophonetic knowledge (see Docherty & Mendoza-Denton 2012 for a review).

For example, Pierrehumbert (2006: 516) emphasizes that language is a collective behavior and that people "match their language systems to each other, and group themselves into social networks of people who share the same language." Furthermore, social expectations have a profound effect on speech perception (as outlined by, e.g., Warren & Hay 2012). In addition, it is evident that lexical and socio-indexical knowledge develop in tandem. In fact, Foulkes (2010) argues that social context provides the earliest consistent categories that a child perceives, allowing him to begin to organize the phonetic variation he hears into categories, as for example, by noting correlations between acoustic features (such as f0) and differences between important persons in the environment (such as female and male adult caretakers). Such evidence highlights the fact that both linguistic and social information are learned, and that social information may even facilitate linguistic learning.

It is clear that understanding social-indexical variation is critical to understanding the core nature of linguistic units, how they are perceived, and how they change. Recent research has documented a variety of ways in which social information attributed to a speaker influences perception of speech sounds produced by that speaker (Johnson et al. 1999; Hay et al. 2006). Furthermore, a variety of evidence indicates that perception and production change based on exposure to variable input within lexically licensed phonological forms, including phonological category plasticity with natural exposure (Evans & Iverson 2004), perceptual learning in the lab (e.g., Norris et al. 2003; Kraljic et al. 2008), and even in passive imitation (Delvaux & Soquet 2007). Understanding category change can in turn help us understand language change over time at the macro and micro levels, from sound changes to phonetic accommodation and entrainment between partners in conversation (e.g., Pardo 2006; Babel 2009; Pierrehumbert 2012). Socio-indexical variation, then, is an additional frontier which challenges our assumptions but ultimately provides a broader and richer foundation for improving our understanding of linguistic knowledge. In many ways, social-indexical variation is parallel to sentence-level prosodic information such as intonation in being clearly governed by grammatical principles, but serving to express both linguistic and non-linguistic information. In the case of intonation, the form of the intonational sequence is constrained by the phonological and syntactic content of the utterance; and in reflecting it, supports its perception (highlighting phrasal heads, edges, and hierarchical structure) while also helping to convey non-lexical information such as emphasis, expectation, and turn-taking in discourse. While mainstream phonologists may not have viewed intonation as part of the phonology in the days of *SPE*, there are few if any now who would deny its fundamentally linguistic nature. We argue that in time sociophonetic variation will be understood in the same way.

18.4.5 Multimodal information in speech processing

Another major thread of research that benefits from the broader and more inclusive research perspective of LP is an increased understanding of the role of multimodal information in speech processing. Work on signed languages has made abundantly clear the potential for different perceptual channels to be employed for human communication and has raised interesting questions about linguistic representations and the extent to which phonology is independent of, or depends on, the "machine/apparatus" used in producing and perceiving language (see Chapter 15 Substance free phonology and Chapter 16 The phonology of sign languages, this volume). Here we briefly discuss a different sense of multimodality, focusing on the role of vision in speech perception. The practice of lip-reading demonstrates that segmental information is available to some extent in the face of the speaker. Improved technology has made possible finer analysis and modeling of the segmental information in the face, as for example in Jiang et al. (2007) who report the correlation of 3-D facial movement data and identification of consonants in CV syllables. That we normally integrate visual and auditory information is also intuitive. As demonstrated strikingly in work on what has been referred to as the McGurk effect (MacDonald & McGurk 1976), this integration can affect the perception of speech segments when there is incongruity between inputs from different modalities. Recent work has extended examination of visual effects to include the role of facial expression and movement in the perception of prosody. For example, Scarborough et al. (2009) demonstrate that listeners can identify the location of phrasal stress (and to a lesser extent, lexical stress) at above-chance levels solely on the basis of information in the face of the speaker, especially chin position. Swerts & Krahmer (2008) demonstrate that listeners identify phrasal stress more quickly when facial and auditory cues are congruent and that this effect holds near phrase edges but not in medial position. The effect of visual information on processing of prosody extends even to syntactic parsing. Borràs-Comes et al. (2014) find that Catalan and Dutch speakers can identify whether a phrase is a question or a statement, above chance level, with facial information only. Visual information can also influence perception by providing relevant contextual information about the acoustic structure of speech. For example, Kraljic et al. (2008) show that an acoustically ambiguous sound will not affect phoneme boundaries when it can be attributed to an incidental consequence of the speech situation (e.g., a pen in the mouth of the speaker) but the same ambiguity causes a shift in phoneme boundaries when no such visual attribution is available.

In this necessarily brief review, we have highlighted some of the ways that LP has addressed Ohala's (2007) first two questions mentioned in section 18.3, and we hope that we have shown how LP has indeed changed the way we do phonology. Equally rich are LP's contributions to a deeper understanding of language acquisition and learning, how language changes over time and why common patterns are seen across diverse languages, and how sound is associated with meaning as part of both a representational and motoric system.

18.5 Methods

One of the most profound changes brought about by the LP community is the use of a diverse set of methods of gaining information about language knowledge and use. The effect of a multidisciplinary community of phonologists, phoneticians, and researchers in allied fields is that the sum of the methods defines an extensive toolkit. More researchers are using these tools, and increasingly researchers are combining methodologies in innovative ways. Here we provide an overview of methodological perspectives as well as some specific types of data collection and analysis used by laboratory phonologists. We aim to cover some of the most well-established and regularly used methods as well as some of the newer approaches which are likely to be commonly used in future work.

18.5.1 Data collection tools, methods, and paradigms

The LP community is committed to both sharing of existing methodologies and development of new ones, as well as the integration of multiple methods. Among the methods of gathering behavioral data, we can identify the common analysis methods of phonology and phonetics from the last century, as well as a wealth of new techniques that probe the core questions of phonology (some of which were discussed in section 18.4).

A longstanding source of data for phonological analysis is impressionistic transcription and introspective judgment. While this tool has formed the basis of insightful work for decades, the fact is that it provides data that are truly the tip of the iceberg of what we know and can understand about speech. Many of the experimental techniques now being adopted by laboratory phonologists, including analysis of speaker judgments and production by multiple speakers, are only slightly more technically involved while offering a richer picture.

Acoustic analysis of audio data is a key tool used to gain insight into aspects of speech production and properties relevant to perception. Acoustic analysis is now frequently done with Praat (Boersma & Weenink 2014), which established a new standard for speech analysis as a free platform, easily accessible to anyone with a computer and internet access. Acoustic properties studied are too numerous to cover comprehensively, ranging from acoustic durations, to pitch, voice quality, vowel quality, and consonant features. Turk et al. (2006) is a helpful resource on methodological issues relating to determining segment durations from audio data.

Aerodynamic and articulatory data have a long history in phonetics. Aerodynamic data include oral, nasal, and sublaryngeal air pressure as well as oral and nasal airflow. In addition to providing direct information about the aerodynamic conditions necessary for production of particular sounds such as fricatives and voiced sounds (see Shadle 2012; Hanson 2012), aerodynamic data can also be used to infer positions of articulators as they affect the size of the oral cavity (e.g., Ladefoged & Maddieson 1996), the aperture of the velopharyngeal port (e.g., Krakow & Huffman 1993; Delvaux et al. 2008), and the mode of vibration of the vocal folds (Hanson et al. 2001; Hanson 2012). Articulatory data have come from video data as well as various techniques that track position and/or movement of the articulators, from X-ray to various pellet-tracking systems. The invasiveness and heavy data-processing requirements of many techniques have kept subject numbers and diversity limited for articulatory studies, and post processing of the data is still very labor intensive. With the advent of portable ultrasound systems, and increased access to MRI facilities, we expect there will be more people gathering this kind of data, an increase in the knowledge base which will benefit the field in a variety of ways. Papers with helpful overviews and methods descriptions include Byrd et al. (2009) and Davidson (2012).

Studies of speech perception probe the detection, distinction, and identification of phonemes as well as the process of using auditory input to identify words. MATLAB, Praat, and E-Prime are common platforms for perceptual studies in a laboratory setting. Behavioral studies of speech perception regularly employ measures of categorization and perceptual sensitivity (see overviews in Holt & Lotto 2010; Iverson 2012). A variety of tasks have probed the interplay between lexical and phonological information in speech perception (see discussion in Frisch et al. 2000). Lexical decision tasks and word recognition tasks are a common method of invoking lexical representations (and the steps in the speech perception process that precede lexical access), with accuracy and reaction times providing a way to probe and quantify effects of lexical structure such as frequency, paradigmatic relations, or neighborhood density (e.g., Ganong 1980; Vitevitch et al. 1997; Watson & Munson 2007). Lexical decision is employed within the perceptual learning paradigm to probe how lexical knowledge influences listeners' adaptation to variable or atypical speech. These studies explore the flexibility of speech categories and the linguistic and social factors which influence how this adaptation operates (e.g., Norris et al. 2003; Kraljic & Samuel 2005).

Another approach to assessing the flexibility of speech categories is the use of imitation and shadowing tasks. Both can be used to test the phonetic flexibility of speakers, and imitation accuracy and speed can be used to assess which phonetic and phonological details are included in speech-planning representations. Priming paradigms test whether prior direct (form priming) or indirect (semantic priming) exposure to a form affects activation levels for a particular word or associated words. Comparing reaction time effects across items in priming paradigms can be used to probe the degree of phonetic detail present in lexical representations (see discussions in Sumner & Samuel 2007; Ernestus & Baayen 2007; Schiller 2012; Albright 2012). Another innovative approach to investigating the relation between phonological and lexical information is demonstrated by Ali & Ingleby (2012), who use differences in rate of McGurk effect-inspired misperceptions to explore listener knowledge of word-internal morphological structure.

Learning of artificial languages exploits properties of a carefully constructed set of nonwords or words presented as representative of an unknown language to see what generalizations learners make beyond the input data (e.g., Peperkamp & Dupoux 2007; see Moreton & Pater 2012 for a recent review). This technique is often used to test for evidence of the effect of phonological universals or markedness preferences on the behavior of learners/listeners when their own language (and its phonological biases) is not being explicitly invoked by the task. In addition to work on the relationship between lexical and phonological representations, work in LP has also considered the influence of semantic and syntactic context on word recognition (e.g., van de Ven et al. 2012).

Additional instrumental methods for studying language processing and representations underlying processing are gaining ground in the field. These include eye-tracking (see Speer 2012 for an overview) and a variety of neurophysiological data such as EEG, MEG, fMRI, and PET (see Idsardi & Poeppel 2012 for an overview). While analysis of these types of data is particularly complex, one important advantage that they can offer is that response to different stimuli can be measured for even the early prelexical aspects of the recognition process. In addition, they can give us information about which differences subjects can detect, independently from the biases that may be expressed in the traditional behavioral identification or discrimination tasks. See for example Scharinger et al. (2011) on EEG evidence of phonological representation of vowel harmony in Turkish.

18.5.2 The move toward larger and more diverse datasets

Laboratory phonologists have been part of a major shift toward use of increasingly larger sets of data and toward more natural and diverse speech samples than the controlled, read speech samples traditionally used. The larger datasets include corpora collected with the goal of providing a planned, long-term resource for use by multiple researchers, and larger datasets made possible by broader subject-sampling strategies, such as contacting subjects and running experiments over the internet.

The use of larger sets of data has been facilitated by the collection of corpora of speech, some of which are annotated with a phonemic or allophonic representation. Contents range widely, from scripted reading to open-ended or task-oriented conversation; from supreme court hearings to vocalizations by children acquiring language. The Audio British National Corpus (www.phon.ox.ac.uk/AudioBNC) is an example of a frequently used resource with a large collection of spoken British English. The Linguistic Data Consortium is one well-known centralized source of a variety of corpora including a variety of languages (www.ldc.upenn.edu/). See also, for example, the list of resources available from the University of Essex (www.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/corpora/list/index2.html). With data storage relatively plentiful and cheap, the wide availability of good quality recording equipment, and the easy use of automatic (pre)processing tools for speech alignment, there will be more and more data of a wide range of types available for study. Cole & Hasegawa-Johnson (2012) provide a brief review of corpora and online resources, and Harrington (2010) provides a good introduction to the analysis of speech corpora. We also note a trend toward development of tools to search or filter data within large corpora (e.g., work by Howell & Rooth http://msuweb.montclair.edu/~howellj/Resources.html). There are also databases focused specifically on segment inventories or surveys of specific phonological structures, such as the UCLA Phonological Segment Inventory Database, LAPSyD (www.lapsyd.ddl.ishlyon.cnrs.fr/), and PBase version 3 (http://137.122.133.199/cgi-bin/pbase3/search.cgi).

Furthermore, with more and more people around the globe being able to use the internet, web-based research has serious potential to increase the size and diversity of our subject samples for studies of topics such as speech production, speech perception, and even language acquisition (see Loehr & Van Guilder 2012 for an overview). These activities include running fairly traditional experiments, but with subjects participating remotely, and large internet surveys, such as text searches (e.g., Zuraw 2006) or more direct "crowd sourcing" or surveys of linguistic usage and intuitions made possible via the internet.

The inclusion of a broader range of speech types has also been one of the hallmarks of work done in LP. For example, more researchers are adopting the approach taken in work in sociophonetics, which has advanced the study of speech with careful attention to the wide range of speech varieties that exist due to stylistic differences and the social groupings signaled by pronunciation. Scobbie & Stuart-Smith (2012) provide recommendations on how to elicit and use more socio-linguistically diverse data in speech research, and Warner (2012) discusses experimental design factors for studying spontaneous speech. Less traditional speech content, such as tongue twisters and speech errors, have begun to be studied more rigorously (Choe & Redford 2012; Frisch & Wright 2002; Goldstein et al. 2007), and imitation has also gained use for testing the processing and representation of sound structure (e.g., Babel 2012). Another area that has shown considerable growth is the study of dialogue and especially interactive task completion, wherein partners collaborate to accomplish a task in which they have similar but non-identical information on a map (e.g., the Map Task; Anderson et al. 1991) or a gameboard (e.g., Speer et al. 2011; Kim et al. 2011). In addition, speech research now regularly includes analysis of foreign-accented and learner speech (see Davidson 2011 for a review).

18.5.3 Data analysis

Statistical analysis is a research tool that has been in longstanding use in phonetics and allied fields. One of the benefits of the multidisciplinarity of LP is that statistical analysis and a variety of data modeling methods are being applied to a broader range of data and questions about representation and processing of speech. MATLAB (www.mathworks.com/) is often used for statistical analysis, as is R. Baayen (2008) describes the use of R for statistical analysis and modeling; see also Gries (2013). Kingston (2012) provides a tutorial introduction to many important statistical concepts critical to work in LP. Clopper (2012) provides an in-depth discussion of clustering and classification methods, with an overview of some of the questions these methods have been applied to in LP type work. A concise introduction to mixed effects models is given in Baayen (2012). See also Sharon Goldwater's Reading list on Bayesian modeling of language (http://homepages.inf.ed.ac.uk/sgwater/reading_list.html). Johnson (2008) is also a good general reference on quantitative methods.

18.6 Future directions

One striking theme running through our survey of LP contributions to the understanding of speech is the tremendous richness of linguistic knowledge that humans continually manipulate and which can change over the course of a lifetime. We are in the relatively early stages of developing accounts of representation and processing that can accommodate the observed data with insight and predictive power. We outline here some of the conceptual and methodological trends that we see as likely to be influential in the coming years of work in the LP framework.

As noted above, LP was initiated during a time of intense concern about the boundaries between the grammar and other, presumably non-linguistic, aspects of human behavior. Decades of careful documentation of differences in fine phonetic details between language groups have called into question common views of the boundaries of phonology, in particular the distinction between the grammar and the strictly physical/mechanical aspects of speech processing. The broad perspective and large and diverse datasets assumed and encouraged within LP promise to continue to move this boundary to include not only fine phonetic differences between languages, but also the patterned variation that arises between socially defined subgroups of a broader language community. We use language in social interaction, and social interactions define objects (groups, actions associated with a group) and meaning (degree of affiliation, role in the group) that require linguistic expression. Through the process of developing accounts for this additional type of variation, we will be moving toward better and more insightful theories of knowledge of sound. We see the integration of social aspects of language as one of the major theoretical challenges that LP will play a major role in addressing.

More generally, as sketched in sections 18.4 and 18.5, LP engages the interplay between the questions of phonology broadly defined and the use of innovative and increasingly multidisciplinary methods that will continue to propel LP forward. One methodological trend which we see continuing is the investigation of multiple dimensions in speech research, such as integrated production and perception studies, or integrated acoustic and imaging studies. Integrating experimental phonetic and psycholinguistic methods and, increasingly, neurolinguistics, provides the empirical foundation for richer and more comprehensive models. The explicitness required to implement a model encourages the rigor and attention to data that is exactly in line with the views of laboratory phonologists. Modeling allows us to explicitly test relations between data and theory. For discussion of recent directions in modeling, see for example Boersma (2012) and Reetz (2012). Bayesian modeling is increasingly being used for multifaceted linguistics data. See for example Feldman et al. (2013), who use a Bayesian model to show how feedback from word segmentation might constrain phonetic category learning.

Another methodological trend, also highlighted in section 18.5, is increasing efforts to use more naturalistic data, and larger datasets, leading to much interest in large corpora and data mining (both of which fall under the increasing emphasis on “big” data). Such data pose a number of interesting challenges in terms of computational power, data management, and analysis of the multitude of factors that come together in shaping the output in such data sources. Greater computational and experimental literacy among linguistics and closer collaboration with computer scientists and speech engineers are important steps in addressing these issues.

Ultimately, better integrated studies and more naturalistic data will help address the increasingly complex questions that phonologists are asking. The relationship between lexical representations, phonology, and phonetics, and ways that the social and communicative aspects of speech interact with the linguistic and representational aspects, are two critical areas that will be better understood through integrated solutions. With respect to the former, while there is widespread agreement that there are both more detailed and more abstract aspects of lexical representation, specific models accounting for the full range of empirical findings are still needed. (For recent discussion see Pisoni & Levi 2007; Nguyen 2012; Ernestus 2014.) In terms of the relationship between social, communicative, and linguistic aspects of speech, much remains to be studied as well. For example, considering just the subset of fine phonetic details that correlate with social constructs, we need to know how much "non-linguistic" phonetic detail is included in lexical representations, and how the constant interplay between social and linguistic meaning is coordinated. (See Foulkes 2010; Docherty & Mendoza-Denton 2012 for recent discussion and review.) Equally important is work integrating our understanding of how language is acquired and learned, and how this relates to individual and collective adult systems, including how these change over the life span and from one generation to the next. (See Munson et al. 2012 for review.)

In sum, the experimental innovations which are the hallmark of LP offer new approaches to the core problems of phonology, and this will continue to be the case as we move forward as a community, developing more integrated methods and models to test and explore the complex and rich questions of the nature of speech and language.

18.7 Further reading

For further reading, we recommend:

1 *Laboratory Phonology*. The Journal of the Association for Laboratory Phonology, de Gruyter Mouton, www.labphon.org/home/journal.
Since 2010, this journal publishes selected papers presented at LabPhon conferences and independent contributions addressing experimental approaches to speech and language.

2 *Papers in Laboratory Phonology I–10*. Volumes I–6, Cambridge: Cambridge University Press. Volumes 7–10, Berlin: de Gruyter Mouton.
The proceedings volumes from the first 10 LabPhon conferences. Earlier volumes focused primarily on issues about the relationship between phonology and phonetics, while the topic coverage of the later volumes has broadened considerably.

3 Cohn, Abigail C., Cécile Fougeron and Marie K. Huffman, eds 2012. *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press.
This Handbook surveys current research in LP, foundational issues, major contributions, and methodologies.

4 Solé, Maria-Josep, Patrice Beddor and Manjari Ohala, eds 2007. *Experimental Approaches to Phonology*. Oxford: Oxford University Press.
This is a diverse volume, with papers on phonetic explanations of phonological universals, sound change, and lexical contrasts, including innovative proposals for future investigations.

5 Scobbie, J. 2007. Interface and overlap in phonetics and phonology. In: G. Ramchand & C. Reiss (eds) *The Oxford Handbook of Linguistic Interfaces*, pp. 17–52. Oxford: Oxford University Press.
This paper argues that phonetics and phonology are clearly distinct, yet there are true intermediate cases that do not belong to either domain.

18.8 Related topics

Chapter 1 The study of phonology in the 21st century: overview and introduction to *The Routledge Handbook of Phonological Theory*
Chapter 17 Phonology as an emergent system

Chapter 19 Articulatory Phonology
Chapter 20 Exemplar theories in phonology
Chapter 22 Statistical phonology

Notes

1 We thank Anna Bosch and S.J. Hannahs for the invitation to step back and reflect on laboratory phonology. We thank Lisa Davidson, Peggy Renwick, Sam Tilsen, and the Cornell PLab group for questions and comments on an earlier draft. We thank Emma Lantz for help with final editing and formatting.
2 Laver (2000: 32): "A single discipline has the attributes of a common paradigm, a common specialist terminology and a common methodological framework. A subject is the combination of the disciplines and topics relevant to a branch of knowledge delimiting a range of related phenomena and processes."
3 We use the term *laboratory phonology* (LP) to refer to this body of research and *LabPhon* to refer to the conferences.
4 For those not familiar with the story, one version is available at www.jainworld.com/literature/story25.htm — accessed Sept 6, 2014.
5 That is, a detailed segmental transcription using the International Phonetic Alphabet (IPA), widely assumed to capture all the linguistically relevant details.

References

Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Albright, A. (2012) Probing underlying representations. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 134–145.

Ali, A. N. & M. Ingleby (2012) Gradience in morphological decomposability: Evidence from the perception of audiovisually incongruent speech. *Laboratory Phonology*. 1 (2). Pp. 263–282.

Anderson, A., M. Bader, E. Bard, E. Boyle, G. M. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson & R. Weinert (1991) The HCRC map task corpus. *Language and Speech*. 34. Pp. 351–366.

Anderson, S. R. (1985) *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago: University of Chicago Press.

Arvaniti, A. (2012) Segment-to-tone association. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 265–274.

Baayen, R. H. (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. H. (2012) Mixed-effects models. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 668–677.

Babel, M. (2009) *Phonetic and Social Selectivity in Speech Accommodation*. Unpublished Ph.D. Dissertation, Department of Linguistics, Berkeley, CA: University of California.

Babel, M. (2012) Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*. 40 (1). Pp. 177–189.

Barry, M. (1992) Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication*. 11. Pp. 393–400.

Beckman, M. E. & J. Kingston (1990, reprinted 2012) Introduction. In J. Kingston & M. Beckman (eds). *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press. Pp. 1–16. (reprinted in Cohn, Fougeron, and Huffman 2012).

Bell, A., D. Jurafsky, E. Fosler-Lussier, E. Girand, & D. Gildea. (2003) Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*. 113 (2). Pp. 1001–1024.

Blumstein, S. & K. N. Stevens (1981) Phonetic features and acoustic invariance in speech. *Cognition*. 10 (1–3). Pp. 25–32.

Boersma, P. (2012) Modelling phonological category learning. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 207–218.

Boersma, P. & D. Weenink (2014) Praat: Doing phonetics by computer [Computer program]. Version 5.3.80, retrieved 29 June 2014 from www.praat.org/

Borràs-Comes, J., C. C. L. Kaland, P. Prieto & M. G. J. Swerts (2014) Audiovisual correlates of interrogativity: A comparative analysis of Catalan and Dutch. *Journal Nonverbal Behavior*. 38 (1). Pp. 53–66.

Browman, C. P. & L. Goldstein (1990) Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*. 18. Pp. 299–320.

Bürki, A., M. Ernestus, C. Gendrot, C. Fougeron & U. H. Frauenfelder (2011) What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *Journal of the Acoustical Society of America*. 130. Pp. 3980–3991. doi:10.1121/1.3658386

Byrd, D., S. Tobin, E. Bresch & S. Narayanan (2009) Timing effects of syllable structure and stress on nasals: A real-time MRI examination. *Journal of Phonetics*. 37. Pp. 97–110.

Chater, N., A. Clark, A. Perfors & J. Goldsmith (2015) *Empiricist Approaches to Language Acquisition*. Oxford: Oxford University Press.

Cho, T. (2011) Laboratory phonology. In N. C. Kula, B. Botma & K. Nasukawa (eds). *The Continuum Companion to Phonology*. London/New York: Continuum. Pp. 343–368.

Choe, W. K. & M. A. Redford (2012) The relationship between speech errors and prosodic phrase boundaries. *Laboratory Phonology*. 3. Pp. 5–26.

Chomsky, N. (1965) *Aspects of a Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. & M. Halle (1968) *The Sound Pattern of English*. New York: Harper & Row.

Clopper, C. (2012) Clustering and classification methods. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 678–692.

Cohn, A. C. (2010) Laboratory phonology: Past successes and current questions, challenges, and goals. In C. Fougeron, B. Kühnert, M. D'Imperio & N. Vallé (eds). *Laboratory Phonology 10: Variation, Phonetic Detail and Phonological Representation*. Berlin: Mouton de Gruyter. Pp. 3–29.

Cohn, A. C., C. Fougeron & M. Huffman (eds) (2012) *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press.

Cohn, A. C. & M. Huffman (2013) Interface between phonology and phonetics. In M. Aronoff (ed). *Oxford Bibliographies Online*. www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0168.xml.

Cole, J. & M. Hasegawa-Johnson (2012) Corpus phonology with speech resources. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 431–440.

Coleman, J. S. (2002) Phonetic representations in the mental lexicon. In J. Durand & B. Laks (eds). *Phonetics, Phonology, and Cognition*. Oxford: Oxford University Press. Pp. 96–130.

Coleman, J. S. (2003) Discovering the acoustic correlates of phonological contrasts. *Journal of Phonetics*. 31. Pp. 351–372.

Conference Proceedings: From Sound to Sense: 50+ Years of Discoveries in Speech Communication (2004) www.rle.mit.edu/soundtosense/conference/pages/invited.htm.

Croot, K. (2010) The emergent paradigm in laboratory phonology: Phonological categories and statistical generalization. *Laboratory Phonology*. 11 (2). Pp. 415–424.

Davidson, L. (2011) Phonetic and phonological factors in the second language production of phonemes and phonotactics. *Language and Linguistics Compass*. 5 (3). Pp. 126–139.

Davidson, L. (2012) Ultrasound as a tool for speech research. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 484–495.

Delvaux, V., D. Didier, B. Harmegnies & A. Soquet (2008) The aerodynamics of nasalization in French. *Journal of Phonetics*. 36 (4). Pp. 578–606.

Delvaux, V. & A. Soquet (2007) The influence of ambient speech on adult speech production through unintentional imitation. *Phonetica*. 64. Pp. 145–173.

Dilley, L., S. Shattuck-Hufnagel & M. Ostendorf (1996) Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*. 24. Pp. 423–444.

D'Imperio, M. (2012) Prosodic representations. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 275–287.

Docherty, G. & N. Mendoza-Denton (2012) Speaker-related variation – sociophonetic factors. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 43–60.

Ellis, L. & W. J. Hardcastle (2002) Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*. 30. Pp. 373–396.

Ernestus, M. (2012) Message related variation: Segmental within speaker variation. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 93–102.

Ernestus, M. (2014) Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*. 142. Pp. 27–41.

Ernestus, M. & R. H. Baayen (2007) Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes*. 22 (1). Pp. 1–24.

Esteve-Gibert, N., F. Pons, L. Bosch & P. Priet (2014) Are gesture and prosodic prominences always coordinated? Evidence from perception and production. In *Proceedings of Speech Prosody 2014*. Dublin, Ireland. Pp. 222–226.

Esteve-Gibert, N. & P. Prieto (2013) Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research*. 56 (8). Pp. 850–864.

Evans, B. G. & P. Iverson (2004) Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America*. 115. Pp. 352–361.

Feldman, N. H., T. L. Griffiths, S. Goldwater & J. L. Morgan (2013) A role for the developing lexicon in phonetic category acquisition. *Psychological Review*. 120 (4). Pp. 751–778.

Fougeron, C. & P. Keating (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*. 101. Pp. 3728–3740.

Foulkes, P. (2010) Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology*. 1 (1). Pp. 5–39.

Foulkes, P. & G. Docherty (2006) The social life of phonetics and phonology. *Journal of Phonetics*. 34 (4). Pp. 409–438.

Frisch, S. A., N. Large & D. Pisoni (2000) Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*. 42. Pp. 481–496.

Frisch, S. A. & R. Wright (2002) The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics*. 30. Pp. 139–162.

Gafos, D. & L. Goldstein (2012) Articulatory representation and organization. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 220–231.

Gahl, S. (2015) Lexical competition in vowel articulation revisited: Vowel dispersion in the East/Hard database. *Journal of Phonetics*. 49. Pp. 96–116.

Ganong, W. F. (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*. 6. Pp. 110–125.

Goldsmith, J. (1976) *Autosegmental phonology*. PhD dissertation, MIT. [Published by New York: Garland Press, 1980.]

Goldsmith, J. (1995) *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell.

Goldsmith, J. & B. Laks (to appear) Generative phonology: Its origins, its principles, and its successors. In L. Waugh, J. E. Joseph, & M. M. Burston (eds). *The Cambridge History of Linguistics*. Cambridge: Cambridge University Press.

Goldstein, L., M. Pouplier, L. Chen, E. Saltzman & D. Byrd (2007) Dynamic action units slip in speech production errors. *Cognition*. 103. Pp. 386–412.

Gow, D. W. (2001) Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*. 45. Pp. 133–159.

Gow, D. W. & B. McMurray (2004) From sound to sense and back again: The integration of lexical and speech processes. *The Proceedings of From Sound to Sense: 50+ years of discoveries in Speech Communication*. MIT, Cambridge, MA. Pp. 118–132.

Gries, S. Th. (2013) *Statistics for Linguistics with R: A Practical Introduction*. Berlin & New York: De Gruyter Mouton.

Halle, M. (1959) *The Sound Pattern of Russian*. The Hague: Mouton.

Hanson, H. M. (2012) Methodologies used to investigate laryngeal function and aerodynamic properties of speech. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 496–510.

Hanson, H. M., K. N. Stevens, H.-K. Kuo, M. Y. Chen & J. Slifka (2001) Towards models of phonation. *Journal of Phonetics*. 29. Pp. 451–480.

Harrington, J. (2010) *The Phonetic Analysis of Speech Corpora*. Oxford: Wiley-Blackwell.

Hawkins, S. (2012) The lexicon: Not just elusive, but illusory? In A. Cohn, C. Fougeron, & M. Huffman (eds). *Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 162–171.

Hawkins, S., & N. Nguyen (2004) Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics*. 32. Pp. 199–231.

Hay, J., P. Warren & K. Drager (2006) Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*. 34 (4). Pp. 458–484.

Holt, L. L. & A. J. Lotto (2010) Speech perception as categorization. *Attention, Perception & Psychophysics*. 72. Pp. 1218–1227.

Hoole, P., B. Kühnert & M. Pouplier (2012) System related variation. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 115–132.

Idsardi, W. J. & Poeppel, D. (2012) Neurophysiological techniques in laboratory phonology. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 593–605.

Iverson, P. (2012) Measuring phonetic perception in adults. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 572–579.

Jiang, J., A. Alwan, P. Keating, E. Auer & L. Bernstein (2007) Similarity structure in visual speech perception and optical phonetic signals. *Perception and Psychophysics*. 69 (7). Pp. 1070–1083.

Johnson, K. (1997) Speech perception without speaker normalization. In K. Johnson & J. Mullennix (eds). *Talker Variability in Speech Processing*. San Diego: Academic Press.

Johnson, K. (2008) *Quantitative Methods in Linguistics*. Oxford: Blackwell.

Johnson, K., E. A. Strand & M. D'Imperio (1999) Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*. 27. Pp. 359–384.

Jurafsky, D. (2003) Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (eds). *Probabilistic Linguistics*. Cambridge, MA: The MIT Press. Pp. 39–95.

Keating, P. (1985) Universal phonetics and the organization of grammars. In V. Fromkin (ed). *Phonetic Linguistics*. San Diego: Academic Press. Pp. 115–132.

Kelly, J. & J. Local (1989) *Doing Phonology: Observing, Recording, Interpreting*. Manchester University Press.

Kim, M., W. S. Horton & A. R. Bradlow (2011) Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Journal of Laboratory Phonology*. 2. Pp. 125–156.

Kingston, J. (2012) Statistical methods in laboratory phonology. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 644–667.

Kohler, K. (2000) The future of phonetics. *Journal of the International Phonetic Association*. 30. Pp. 1–24.
Krakow, R. & M. Huffman (1993) Instruments and techniques for investigating nasalization and velopharyngeal function in the laboratory: An introduction. In M. Huffman & R. Krakow (eds). *Nasals, Nasalization and the Velum*. San Diego: Academic Press. Pp. 3–59.
Kraljic, T. & A. G. Samuel (2005) Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*. 51. Pp. 141–178.
Kraljic, T., A. G. Samuel & S. E. Brennan (2008) First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*. 19. Pp. 332–338.
Ladd, D. R. (2006) Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Italian Journal of Linguistics*. 18 (1). Pp. 19–38.
Ladd, D. R. (2008) *Intonational Phonology*. 2nd Ed. Cambridge, UK: Cambridge University Press.
Ladd, D. R. (2011) Phonetics in phonology. In J. Goldsmith, J. Riggle, & A. Yu (eds). *Handbook of Phonological Theory*. 2nd Ed. Oxford: Blackwell. Pp. 348–373.
Ladd, D. R., D. Faulkner, H. Faulkner & A. Schepman (1999) Constant "Segmental anchoring" of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America*. 106 (3). Pp. 1543–1554.
Ladd, D. R., I. Mennen & A. Schepman (2000) Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*. 107 (5). Pp. 2685–2695.
Ladefoged, P. & I. Maddieson (1996) *The Sounds of the World's Languages*. Oxford, UK: Blackwell.
Laver, J. (2000) The nature of phonetics. *Journal of the International Phonetic Association*. 2000 (30). Pp. 31–36.
Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle & A. Marchal (eds). *Speech production and speech modeling*. Dordrecht, The Netherlands: Kluwer Academic Publishers. Pp. 403–439.
Local, J. (2003) Variable domains and variable relevance: Interpreting phonetic exponents. *Journal of Phonetics*. 31. Pp. 321–339.
Loehr, D. (2012) Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*. 3 (1). Pp. 71–89.
Loehr, D. & L. Van Guilder (2012) Using the internet for collecting phonological data. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 441–449.
MacDonald, J. & H. McGurk (1976) Hearing lips and seeing voices. *Nature*. 264 December 23/30. Pp. 746–774.
Moreton, E. & J. Pater (2012). Structure and substance in artificial-phonology learning: Part I, structure: Part II, substance. *Language and Linguistics Compass*. 6 (11). Pp. 686–701 and 702–718.
Munson, B., J. Edwards & M. E. Beckman (2012) Phonological representations in language acquisition: Climbing the ladder of abstraction. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 288–309.
Nguyen, N. (2012) Representations of speech sound patterns in the speaker's brain: Insights from perception studies. In A. Cohn, C. Fougeron, & M. Huffman (eds). *Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 359–368.
Nolan, F. (1992) The descriptive role of segments: Evidence from assimilation. In G. J. Docherty & D. R. Ladd (eds). *Gesture, Segment, Prosody: Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press. Pp. 261–289.
Norris, D., J. M. McQueen & A. Cutler (2003) Perceptual learning in speech. *Cognitive Psychology*. 47. Pp. 204–238.
Ohala, J. J. (2000) Phonetics in the free market of scientific ideas and results. *Journal of the International Phonetic Association*. 30. Pp. 25–29.
Ohala, J. J. (2007) Methods in phonology. In M. J. Solé, P. S. Beddor, & M. Ohala (eds). *Experimental Approaches to Phonology*. Oxford: Oxford University Press. Pp. 2–6.
Ohala, J. J. & J. J. Jaeger (eds) (1986) *Experimental Phonology*. Orlando, FL: Academic Press.

Pardo, J. S. (2006) On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*. 119 (4). Pp. 2382–2393.

Peperkamp, S. & E. Dupoux (2007) Learning the mapping from surface to underlying representations in an artificial language. In J. Cole & J. Hualde (eds). *Laboratory Phonology 9*. Berlin: Mouton de Gruyter. Pp. 315–338.

Pierrehumbert, J. B. (1980) *The Phonology and Phonetics of English Intonation*. Unpublished Ph.D. dissertation. Cambridge, MA: Massachusetts Institute of Technology.

Pierrehumbert, J. B. (2001) Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (eds). *Frequency Effects and the Emergence of Lexical Structure*. Amsterdam: John Benjamins. Pp. 137–157.

Pierrehumbert, J. B. (2006) The next toolkit. *Journal of Phonetics*. 34. Pp. 516–530.

Pierrehumbert, J. B. (2012) The dynamic lexicon. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 173–183.

Pierrehumbert, J. B. & M. E. Beckman (1988) *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Pierrehumbert, J. B., M. E. Beckman & D. R. Ladd (2000) Conceptual foundations of phonology as a laboratory Science. In N. Burton-Roberts, P. Carr, & G. Docherty (eds). *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press (reprinted in Cohn, Fougeron, and Huffman 2012). Pp. 273–304.

Pierrehumbert, J. B. & C. Clopper (2010) What is LabPhon? And where is it going? In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallé (eds). *Laboratory Phonology 10: Variation, Phonetic Detail and Phonological Representation*. Berlin: Mouton de Gruyter. Pp. 113–132.

Pierrehumbert, J. B. & D. Talkin (1992) Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (eds). *Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press. Pp. 90–117.

Pisoni, D. B. & S. V. Levi (2007). Representations and representational specificity in speech perception and spoken word recognition. In G. Gaskell (ed). *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press. Pp. 3–18.

Reetz, H. (2012) Speech manipulation, synthesis, and automatic speech recognition in laboratory phonology. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 450–457.

Repp, B. H. (1984) Categorical perception: Issues, methods and findings. In N. Lass (ed). *Speech and Language (Vol. 10): Advances in Basic Research and Practice*. Orlando: Academic Press. Pp. 244–322.

Scarborough, R., P. Keating, S. Mattys, T. Cho & A. Alwan (2009) Optical phonetics and visual perception of lexical and phrasal stress in English. *Language and Speech*. 51 (2/3). Pp. 135–175.

Scharinger, M., S. Poe & W. Idsardi (2011) Neuromagnetic reflections of harmony and constraint violations in Turkish. *Laboratory Phonology*. 2(1). Pp. 99–124.

Schiller, N. O. (2012) Experimental methods and designs to investigate phonological encoding of spoken language. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 562–575.

Scobbie, J. (2007) Interface and overlap in phonetics and phonology. In G. Ramchand & C. Reiss (eds). *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press. Pp. 17–52.

Scobbie, J. & J. Stuart-Smith (2012) Socially stratified sampling in laboratory-based phonological experimentation. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 607–620.

Shadle, C. H. (2012) The acoustics and aerodynamics of fricatives. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 511–526.

Silverman, D. (2011) Usage-based phonology. In N. Kula, B. Botma, & K. Nasukawa (eds). *The Continuum Companion to Phonology*. London/New York: Continuum. Pp. 369–394.

Solé, M.-J., P. S. Beddor & M. Ohala (2007) *Experimental Approaches to Phonology*. Oxford: Oxford University Press.

Speer, S. R. (2012) Eye movements as a dependent measure in research on spoken language. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 580–592.

Speer, S. R., P. Warren & A. J. Schafer (2011) Situationally independent prosodic phrasing. *Laboratory Phonology*. 2. Pp. 35–98.

Sumner, M. & A. G. Samuel (2005) Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language*. 52. Pp. 322–338.

Sumner, M. & A. G. Samuel (2007) Lexical inhibition and sublexical facilitation are surprisingly long lasting. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 33. Pp. 769–790.

Swerts, M. G. J. & E. J. Krahmer (2008) Facial expressions and prosodic prominence: Effects of modality and facial area. *Journal of Phonetics*. 36 (2). Pp. 219–238.

Turk, A., S. Nakai & M. Sugahara (2006) Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, & P. Augurzky (eds). *Methods in Empirical Prosody Research*. Berlin: Mouton de Gruyter. Pp. 1–28.

Van de Ven, M., M. Ernestus & R. Schreuder (2012) Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Laboratory Phonology*. 3. Pp. 455–481.

Vitevitch, M. S., P. A. Luce, J. Charles-Luce & D. Kemmerer (1997) Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*. 40. Pp. 47–62.

Wagner, P., Z. Malisz & S. Kopp (2014) Gesture and speech in interaction: An overview. *Speech Communication*. 57. Pp. 209–232.

Warner, N. (2012) Methods for studying spontaneous speech. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 621–633.

Warner, N., A. Jongman, J. Sereno & R. Kemps (2004) Incomplete neutralization and other subphonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*. 32. Pp. 251–276.

Warren, P. & J. Hay (2012) Methods and experimental design for studying sociophonetic variation. In A. Cohn, C. Fougeron, & M. Huffman (eds). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press. Pp. 634–642.

Watson, P. & B. Munson (2007) The influence of phonological neighborhood density and word frequency on vowel-space dispersion in older and younger adults. In J. Trouvain & W. Barry (eds). *Proceedings of the International Congress on Phonetic Sciences*. Saarbrücken, Germany: University of Saarland. Pp. 561–564.

West, P. (1999) Perception of distributed coarticulatory properties in English /l/ and /ɹ/. *Journal of Phonetics*. 27. Pp. 405–426.

Wright, R. A. (2004) A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (eds). *Phonetically Based Phonology*. Cambridge: Cambridge University Press. Pp. 34–57.

Zsiga, E. (1995) An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In B. Connell & A. Arvaniti (eds). *Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press. Pp. 282–302.

Zuraw, K. (2006) Using the web as a phonological corpus: A case study from Tagalog. *EACL2006: Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics/Proceedings of the 2nd International Workshop on Web as Corpus*. Trento, Italy.

19
Articulatory Phonology

Nancy Hall

19.1 Introduction

The central premise of Articulatory Phonology (AP) is that the representational units of phonology correspond to speech production events. Whereas most phonological theories assume that speakers mentally represent a word in terms of features or segments, AP uses a very different set of representations: articulatory *gestures*, and the *coordination structure* that determines their relative timing. Gestures act both as units of contrast and as units of speech production, essentially erasing the traditional distinction between phonology and phonetics.

AP has played a large role in the trend towards Laboratory Phonology (see Chapter 18, this volume). Developed in large part at Haskins Laboratories, through the work of Catherine Browman, Louis Goldstein, and colleagues (e.g., Browman & Goldstein 1986, 1988, 1989, 1990, 1992a, 1992b; Byrd 1995, 1996; Nam & Saltzman 2003; Nam et al. 2004, 2006, 2009), the model has been developed and tested through extensive articulatory phonetics research, as well as computational simulations of speech production.

This chapter is laid out as follows: section 19.2 introduces the basic mechanics of the theory, including the gestural representational system, the computational system that produces and interprets gestural representations, and the types of articulatory data that proposed representations are often based on. AP work on speech errors is reviewed as a case study. section 19.3 reviews AP analyses of a variety of phonological processes, both categorical and non-categorical. Section 19.4 gives an overview of AP work on syllable structure, particularly the coupling model, an important recent development which attempts to explain onset–coda asymmetries as results of different gestural coupling relations. Section 19.5 covers several current trends in AP research, including work on modelling phonological acquisition, morphological structure, tone, and intonation.

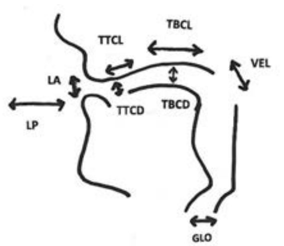
19.2 Gestures as phonological primitives

In AP, the basic units of phonological representation are not features or segments, but articulatory gestures. A gesture can be thought of as a task, a goal to be achieved through articulatory movements. Typical tasks in speech production might include “form a closure with the lips," "spread the vocal folds," or "position the tongue body close to but not touching the velum."

Formally, articulatory goals are defined in terms of "tract variables" (Browman & Goldstein 1989). The most commonly used are those below. Each tract variable refers to a region of the vocal tract. Some of the goals specify a degree of constriction; others specify a location of constriction.

(1) **Tract variables**

lip aperture (LA)	tongue body constriction location (TBCL)
lip protrusion (LP or PRO)	tongue body constriction degree (TBCD)
tongue tip constriction location (TTCL)	velic aperture (VEL)
tongue tip constriction degree (TTCD)	glottal aperture (GLO)



Each of these variables can take a range of values (Browman & Goldstein 1989: 209), as shown below.

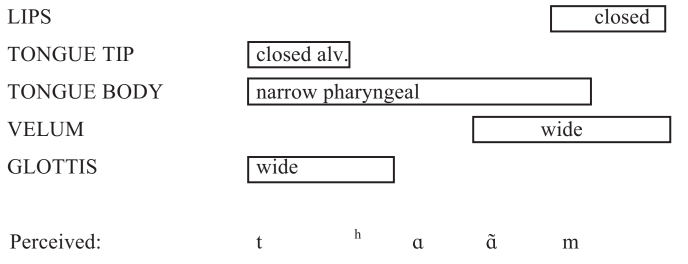
(2) Constriction degree values: closed critical narrow mid wide
Constriction location values: protruded labial dental alveolar postalveolar palatal velar uvular pharyngeal

Segments have no formal role in most AP work; they are regarded as epiphenomenal. Typically, what would be considered a segment in other frameworks corresponds to several gestures in AP. A transcription of [t], for example, would correspond to the gestures "GLO wide" (for voicelessness) and "TT alveolar closed." Since they refer to the same articulator, the variables of TTCD and TTCL must be specified together, as must TBCL and TBCD, and LA and LP.

The set of tract variables will probably expand as more sounds are modelled. For example, Browman & Goldstein (1989: 228) and Proctor (2010: 93) suggest that a tract variable for tongue tip closure orientation (TTCO) would be useful for capturing the difference between apicals and laminals, and a tract variable for cross-sectional shape of the tongue body would help capture the difference between laterals and centrals. Goldstein (1994: 238) notes that a Tongue Root (TR) variable will be needed for gutturals, but that its articulator set has not been explicitly modelled.

Gestures have a duration in time, and overlap with one another. This overlap is represented in a "gestural score." Below is an example of a possible gestural score for the word *Tom* [tʰãm]. (All gestural scores in this chapter should be understood as schematic and not necessarily to scale; some details may be inferred, and others have been simplified to illustrate the points at hand. For precise temporal and spatial data, please refer to the original sources cited.)

(3) Gestural score of [t^hãm] (horizontal axis = time)



Three gestures begin at about the same time: the tongue tip closure and glottal opening of the [t]; and the tongue body constriction of the [ɑ] (note that low vowels are considered pharyngeal constrictions). The idea of ordering is quite different in AP compared to most other theories. In a segmental representation of *Tom*, we would say that [t] comes before [a], but here the gestures associated with [t] simply *end* earlier than the tongue body gesture associated with [ɑ]. The tongue body gesture is itself overlapped by the velum opening gesture, and since the velum lowers before the lips close, the end of the vowel is nasalized.

Note that the glottis is opened for voicelessness (and creates aspiration by staying open after the tongue tip constriction ends). There is no gesture for voicing; a glottal aperture that produces voicing is assumed to be the default state (Browman & Goldstein 1989: 239). The velum is assumed to be raised by default and requires a lowering gesture to produce nasals (as shown above), but once lowered it requires a raising gesture to return to closure. Browman & Goldstein (1986: 242) note that:

> the decision to treat velic opening and closing as two separate gestures, as compared with the glottal and oral gestures that incorporate both opening and closing, is based on the fact that each velic gesture may act as a word-level phenomenon, so that the velum can possibly be held in either a closed or an open position indefinitely.

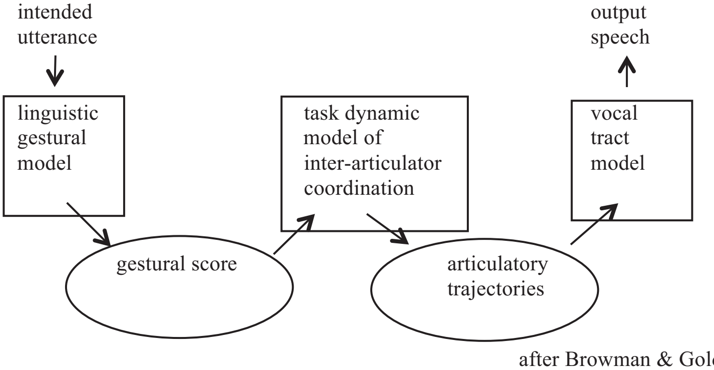
It is important to understand that a gesture, in AP terms, does not refer to the articulatory movements themselves. It specifies a goal, not a means of achieving the goal. The same gesture may cause different movements in different contexts, and possibly even involve different articulators, depending on factors such as what position the articulators start in. For example, imagine how you would achieve the task "tongue tip alveolar closure" (for a [t] or [d], say), starting with the tongue in position for [i], as in "beet." This would involve a small upward movement of the tongue, and likely no movement of the jaw, since the jaw would already be fairly high. The movements involved would be different than if the tongue started in the low back position of [ɑ], as in "bot" – in that case, the tongue would need to move farther and from a different direction, and probably the jaw would raise. Yet the abstract closure gesture would be identical in both cases. Incidentally, if such a gesture was activated when the tongue was already creating an alveolar closure – perhaps for a preceding [n], as in *bent* – then the closure task would actually be accomplished without any movement. It would simply cause the tongue to stay in place longer.

By a similar token, the presence of a gesture does not necessarily mean the gestural target is fully achieved. In fast speech, a closure gesture might not result in a complete stoppage of airflow. Yet again, the abstract gesture would be the same.

19.2.1 Computational modeling of gestures

Gestural scores are part of a larger computational model of speech production, whose basic structure is shown below (Browman & Goldstein 1990: 342). An intended utterance provides input to a linguistic gestural model. This model determines the coordination of gestures required for a particular utterance, and produces a gestural score representing the results.

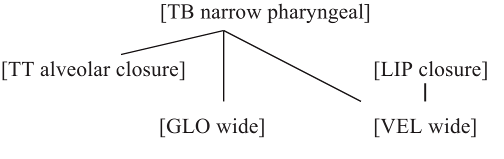
(4)



after Browman & Goldstein (1990:342)

Within the linguistic gestural model, certain pairs of gestures are temporally coordinated with respect to one another. For example, the word [t^{h}ãm] might, hypothetically, start with formal coordination relations between the pairs of gestures linked by lines below:

(5)



Determining which gestures are actually coordinated is an empirical question, answerable only through articulatory studies. Coordinated pairs of gestures are recognized largely through showing relatively stable timing with respect to one another. If a speaker produces a word repeatedly, especially with some variation in speech rate, the relative timing of coordinated gestures (like LIP and VEL in the hypothetical example above) will be more consistent than the relative timing of non-coordinated gestures (like LIP and TB).

Although various details of the model may vary from proposal to proposal, one major shift deserves special mention. Earlier and later AP models assume a very different mechanism for determining the level of overlap between a pair of coordinated gestures. The earlier approach, as sketched in Browman & Goldstein (1992b: 161), was for the linguistic gestural model to specify an alignment between two gestures' internal "landmarks." Landmarks were stages in the articulation of a gesture. This approach was perhaps most fully developed in

Gafos' (2002) analysis of consonant clusters in Moroccan Arabic. In this grammar, each gesture's landmarks included an ONSET, when the articulators first came under active control; a TARGET, when the desired constriction was to be reached; a CENTER, at the midpoint of the constriction; a RELEASE, when movement away from the target began; and an OFFSET, when the articulator ceased to be under active control. When coordinating two gestures, the linguistic gestural model could specify a particular alignment such as CENTER = ONSET (shown below right, where the vertical line indicates the point of alignment).

(6) Internal structure of a gesture Sample alignment CENTER = ONSET

(horizontal axis = time; vertical axis = gesture's activation level)

This approach had the advantage of being computationally explicit, but the disadvantage of allowing too many types of coordination. It has largely been supplanted by the coupled oscillator model (Saltzman & Byrd 2000; Nam & Saltzman 2003; Nam 2007; Nam et al. 2009), which is discussed further in section 4. However, familiarity with the older model is still useful for reading AP literature.

To convert gestural scores to articulatory movements, AP uses a model of task dynamics. Task dynamics is a model of the control of skilled movements, originally developed to describe non-speech movements such as reaching (Bernstein 1967; see Hawkins 1992 for an overview of applications to speech). Based on a gestural score, as well as a parameter for speech rate, the dynamical equations of the task dynamic model determine the actual trajectories that articulators would take. In many cases, overlapping gestures put competing demands on an articulator. In this situation, the task dynamic model "blends" the gestures, creating an articulatory path that compromises between the two goals.

As an implementation of task dynamics, AP researchers generally use TaDA (Task Dynamic Application, available freely at www.haskins.yale.edu/TaDA_download/). Developed by Hosung Nam and colleagues (Nam et al. 2004), one of TaDA's functions is to convert a gestural score to a set of vocal tract shapes. These articulatory trajectories in turn provide input to the vocal tract model HLsyn (Hanson & Stevens 2002), which converts them to an acoustic output.

Using this model, then, a researcher can simulate the articulatory and acoustic output of a hypothetical gestural representation. Such simulations play a large role in AP research. Typically, simulations are run for the purposes of comparison with articulatory (and sometimes acoustic) records of actual speech. A proposed gestural representation is judged as successful to the extent that the simulation and the actual speech match.

19.2.2 Methods: articulatory data collection

Although a gesture is not itself a physical movement, physical movements are the best evidence for inferring gestural structures. For this reason, AP depends heavily on techniques for measuring articulatory motion, such as electromagnetic articulography (EMA/EMMA), ultrasound, X-ray microbeam, and real-time MRI. Even when articulatory records are obtainable, it is not always straightforward to detect the underlying gestural structure, as Gick et al. (2006: 69) comment: "the criteria for determining under what circumstances an observed physical event should be considered phonologically real have been vague in the previous literature on AP (and essentially absent from most other prominent models of phonology)." There is now a body of work aimed at developing algorithms for extracting gestural scores from articulatory speech records (for example, Ramanarayanan et al. 2013).

It is unfortunately much more difficult to infer gestural structures from acoustic records. Although a few articulatory events do have clear acoustic correlates (such as the achievement of a stop closure after a vowel), many do not. For example, it's often impossible to identify the beginning or end of a gesture from the acoustic record, because usually it is masked by other gestures that overlap it. In a word like *tee* [t^hi], the vowel gesture begins at some point during the stop closure, but it's hard to say exactly when. This is a problem, because recent work suggests that the beginnings of gestures are important anchor points in the control of coordination. So although studies do sometimes compare the results of simulations to acoustic records, articulatory records are considered far preferable.

One disadvantage of the reliance on articulatory data is that such data can be expensive and labor intensive to collect and process. For this reason, AP studies tend to be based on a small number of subjects. Languages are quite often described on the basis of a single speaker, and a study with five speakers can be considered large in AP terms. The number of languages described to date is likewise small. Typological studies of dozens of languages, which play an important role in approaches such as Optimality Theory, are impractical in articulatory research. Lack of access to expensive equipment can also be a barrier for new researchers, although it should be noted that some seminal work in AP has been based on existing articulatory corpora rather than purpose-collected data. Several such corpora are now publicly available, such as the X-ray speech database of Munhall et al. (1995).

19.2.3 Case study: articulatory study of speech errors

As an example of how a phenomenon can appear qualitatively different when described from an articulatory as opposed to acoustic viewpoint, consider the recent AP work on speech errors (Pouplier & Goldstein 2005; Goldstein et al. 2007). Most studies of speech errors are based on impressionistic transcriptions, and these studies have converged on some apparently robust generalizations: for example, most speech errors are said to involve moving or substituting segments (as opposed to features), and these errors are said to almost never violate phonotactic rules of the speaker's language. So, for example, *tariffs and barriers* might be mangled to *bariffs and terriers*, but would not become *tbariffs and btarriers*, in a language that does not allow [tb] or [bt] onsets.

However, AP studies such as Goldstein et al. (2007) have shown through EMA that the articulatory reality is different. When speakers were given a tongue-twisting exercise like repeating *top cop* over and over, they would slip and produce tokens that sounded like *cop cop* or *top top*. Yet in the articulatory movement traces, it was evident that the [k]s or [t]s produced in error were not like normal stops. Often, people seemed to produce both gestures at once, with a simultaneous velar and alveolar closure gesture (not necessarily fully achieved). Pouplier & Goldstein (2005) show that such errors are hard to hear correctly; listeners tend to either miss them or hear them as segmental substitutions. Similarly, Goldstein et al. (2007) also found that speech errors might involve only one of the gestures associated with a segment. For example, /m/ has both a bilabial closure gesture and a velar lowering gesture. When speech errors occurred in a phrase like *kim kid*, sometimes only one of the /m/'s gestures would move to the /d/. The velum might lower slightly during the tongue tip gesture, yet without lip movement.

These studies offer a strong challenge to the conclusions of non-articulatory studies: they suggest that many speech errors involve the movement of gestures rather than whole segments, and that the result does not have to conform to the language's phonotactics. Needless to say, this result is also highly consistent with the AP claim that gestures rather than segments are the basic units of speech.

19.3 Representations of phonological phenomena

The gestural representation system of AP means that phonological phenomena such as alternations must also be described in gestural terms. In some cases, these phenomena are analyzed quite differently in AP than in segment and feature-based theories.

19.3.1 Categorical and non-categorical processes

One advantage of gestural representations is their ability to capture the differences between fast and slow speech without fundamentally restructuring the utterance. In theories based on segmental representations, casual speech is often described as characterized by the deletion or substitution of segments (see Browman & Goldstein 1990: 359 for numerous examples). Segments may acoustically disappear: a phrase like *he looked past me* might sound like [hi lʊk pæs mi], with /t/s eliding between two consonants; a word like *support* may sound like [sport], with elision of the schwa. Other segments lenite: an intervocalic /b/ as in *about* might be pronounced as a fricative [β]. Both nasal and oral stops tend to assimilate in place to following stop, so that phrases like *fat cat* may sound like [fækkæt]. In segment-based frameworks, such changes must be analyzed (and transcribed) as categorical changes, governed by rules such as t → Ø/C__C. Yet this flies against evidence, both from articulatory studies and speaker intuitions, that at least some of these changes are actually gradient. The lenited segments resulting from fast-speech assimilation or lenition are not necessarily identical to regular, lexical occurrences of the (apparently) same segments.

Browman & Goldstein (1990) propose that no gestures are deleted in fast speech, nor do gestures change their tract variable values (such as LIP closure). Rather, fast speech causes reductions in the magnitude of gestures and increases in the relative overlap of gestures. Both of these changes can affect the acoustic output, by causing gestures to not reach their targets, or by hiding one gesture behind others. In a famous example, the authors identified a token in a corpus of X-ray films of speech where an English speaker pronounced *perfect memory* with the [t] acoustically absent. Yet the X-ray record showed that the tongue tip gesture was still executed. It was simply inaudible because it was completely overlapped by the closures of the preceding [k] and following [m].

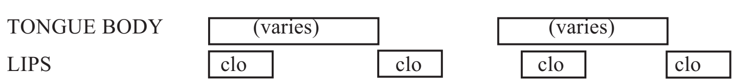
The AP approach does not assume, however, that every assimilation, lenition, etc. is necessarily a result of gestural overlap. Rather, it allows a better description of the difference between categorical and non-categorical changes. For example, Zsiga (1995) compares two processes in which /s/ palatalizes to [ʃ]. In words like *confession*, the (arguable) underlying /s/ that is pronounced in the stem *confess* obligatorily palatalizes, producing [kənfɛʃn̩]. Zsiga shows that this type of derived [ʃ] is indistinguishable from lexical [ʃ]. Yet a different picture emerges for the optional, casual speech phenomenon in which phrases like *press you* are pronounced like [prɛʃu]. Zsiga shows that this [ʃ] is different both from lexical [ʃ] and from the [ʃ] in *confession*. The degree of casual speech palatalization is variable, and some tokens are s-like at the Onset of the fricative, yet ʃ-like by the end. Zsiga proposes that the palatalization in *press you* is caused purely by gestural overlap. When the tongue tip gesture of /s/ overlaps the tongue body gesture of /j/, the blending of the two gestures in the task dynamic model causes the tongue to retract to a more /ʃ/-like position. The palatalization in *confession*, on the other hand, involves some categorical alternation. In Zsiga's model, *confession* involves feature-spreading; her approach is unusual among AP theorists in giving a formal role to features. Another approach would be to assume that *confess* and *confession* underlyingly have different TT gestures.

19.3.2 Presence or absence of gestures: the case of schwa

Another area in which there has been considerable examination of categorical vs. non-categorical gestural changes is vowel alternations, particularly involving schwa. Schwa-vowels present an interesting ambiguity, because it is possible to produce an acoustic schwa without specifying a vocalic tongue body target. This is partly because the tongue position for schwa is similar to the tongue's resting position, to which it returns during periods when it is not under active control.

To see what gestural scores could in principle underlie a schwa, Browman & Goldstein (1992a: 51–54) simulate the production of [pVpəpVp] sequences (where the gestural targets of the Vs vary), using gestural scores like that below. They find that during the gap between the second and third lip closures, where there is no active tongue body gesture, the tongue body dips towards a schwa-like position and a perceptible schwa appears in the acoustic output.

(7) Gestural score of simulated [pVpəpVp]



Similarly, Gick & Wilson (2006) show that the acoustic schwa in words like *fire* [fajəɹ] does not necessarily reflect a schwa-target. The tongue has to pass through a schwa-like configuration on the way from the high front target of /j/ to the low back target of /ɹ/.

It does not appear that all schwas in real speech are targetless, however. Analyzing X-ray data of an English speaker producing sequences like [pipəpipə], Browman & Goldstein (1992a: 51) detect tongue movement towards a possible target associated with the medial schwa, and conclude that the data "argue against the strongest form of the hypothesis that schwa has no tongue target."

Nevertheless, the possibility of producing schwa without active control raises the question of whether targetless schwas occur in natural language. Gafos (2002) argues that such a schwa occurs in final CC clusters in Moroccan Arabic. Words ending in a heterorganic CC cluster have an audible schwa in slow speech but not fast speech: for example, the participle of "write" can be pronounced [kat^{ə}b] or [katb]. Using simulations, Gafos shows that a timing relation of CENTER = ONSET for the oral gestures (such as tongue tip alveolar closure and lip closure in /katb/) produces an audible release at slower rates of speech but no release at higher rates. On the other hand, final clusters of identical consonants have a schwa at all rates of speech, as in [wlas^{ə}s] "swollen gland." Gafos shows that a timing relation of OFFSET = ONSET produces a consistent audible schwa at all speech rates. He argues that this timing relationship reflects a principle of avoiding overlap between identical gestures, a type of gestural Obligatory Contour Principle.

Hall (2006) argues that targetless vowels show different phonological behaviours than vowels that correspond to a tongue body gesture. In a typological study (based on transcriptions), she identifies vowels, described as epenthetic, that have characteristics typical of a targetless vowel. These vowels have qualities that can be explained without positing a distinct gesture (either schwa, or influenced by the qualities of overlapping vowel or consonant gestures); they tend to be optional and disappear at fast speech rates; and they occur in heterorganic clusters, which are more prone to having an acoustic release between the consonants. She argues that vowels with these characteristics also tend to act phonologically invisible: for example, they do not count as a syllable in the stress system or for minimal word requirements; they are ignored in language games; and they fail to trigger phonological processes such as spirantization of a following stop. Speakers may be unaware that the vowels are even present. Furthermore, such vowels tend to occur in CC clusters that are cross-linguistically unmarked, and hence unlikely candidates for phonological repair.

Davidson & Stone (2003) show that targetless schwas may also occur in second language speech. When English speakers are asked to read pseudo-Slavic forms like *zgomu*, they often insert a schwa in a non-native consonant cluster, producing what sounds like [zəgomu]. Yet when these productions are studied by ultrasound, the tongue body position turns out to be different than that in real English words like *succumb* [səkʌm], where a schwa occurs between consonants that have the same TT and TB targets as /zg/. The articulatory trajectory of the tongue in [zəgomu] is consistent with the lack of an articulatory target for the schwa. This suggests that the acoustic schwa may be only a result of low gestural overlap between the consonants, which in turn is probably caused by speakers' lack of experience with coordinating consonant pairs that do not occur in their native language.

These studies illustrate how a common phonological topic like vowel epenthesis is seen differently in AP: the central question is what gestural structure underlies the (acoustic) vowels. The answer to this question may turn out to determine other aspects of the vowels' phonological patterning.

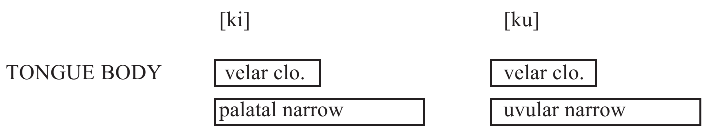
19.3.3 Capturing contrasts and allophony

The topic of contrast is another that is seen differently in AP than in most theories. In AP, contrasts are modelled in terms of gestural specifications, or coordination of gestures; the substance of the contrast is the same as the substance of phonetic realization. There is no precise equivalent for the concepts of *phoneme*, *allophone*, or *feature*.

The methodology of studying contrasts in AP is also non-traditional. For example, Proctor et al. (2010) set out to construct a gestural theory of coronal contrasts in Wubuy. This Australian language has four coronal stops, transcribed as [t ʈ t̠ c]. In traditional, feature-based approaches to describing these contrasts, features would be posited based partly on phonetic descriptions, and partly on phonological patterns such as neutralization in particular environments. In Proctor et al.'s gestural approach, on the other hand, the first step was to collect EMA data of three speakers producing each stop between two vowels. Next, the researchers simulated a hypothetical gestural alignment using TaDA, and compared the results to the EMA data. Where discrepancies were found, in tongue shape or tongue trajectories, the model was iteratively adjusted and re-run to minimize the differences. Phonological distribution patterns played no role in the argumentation concerning the sounds' representation (although distribution does, of course, require explanation in AP). The result of the analysis is not a set of features, but a set of aperture, location, and coordination settings for tongue body and tongue tip constrictions.

What is usually called allophony can result from more than one cause in the AP approach. One cause is gestural blending, as described earlier in the task dynamic model. For example, suppose a velar stop gesture overlaps a vowel gesture. Since both gestures involve the same articulator, the tongue body, the task dynamic model must blend the two. The result is that the actual location of constriction will be different in sequences like [ki] and [ku]: a front vowel will pull the tongue body forward, creating a more forward constriction. The result would usually be described as allophony of the /k/, but at an abstract gestural level, there is no difference between the closure gesture in [ki] and [ku] (Saltzman & Munhall 1989).

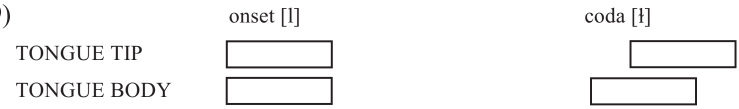
(8)



Of course, languages differ in the extent of CV coarticulation they display. For example, an English velar is only slightly affected by a following vowel, but a Navajo velar is dramatically affected: Navajo /x/ has allophones as divergent as [w] before [o] and [ç] before [i]. Iskarous et al. (2012: 7) propose to model this cross-linguistic variation in coarticulation levels by allowing language-specific settings for blending parameters, which essentially designate the relative strength of conflicting gestural targets. For example, if a language assigns the TBCL of the velar closure gesture a stronger weight than the TBCL of the vowel gesture, then the vowel will have minimal effect on the constriction location of the closure, as in English. If the vowel gesture has a stronger weight, then extreme consonant coarticulation will result, as in Navajo.

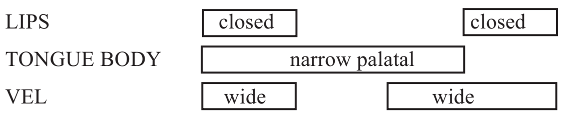
Some allophony, however, is not merely a matter of blending, but reflects differences in the magnitude or coordination of sets of gestures when they occur in different positions within a word or syllable. For example, English /l/ is described as having a "clear" quality in Onset position and a "dark," velarized quality in coda position. This is more than a blending effect; the gestures that correspond to /l/ have a different timing relation in different positions. English /l/ involves both a front tongue tip constriction and a back tongue body constriction. In Onset position, the two gestures begin about simultaneously and the tongue tip gesture is strong; but in coda position, the tongue body gesture precedes the tongue tip gesture and the tongue tip gesture is relatively weak (Sproat & Fujimura 1993; Krakow 1999).

(9)



This turns out to be not an isolated fact about /l/, but a more general pattern in the organization of English syllables. For example, the velic opening and oral closure gestures of a nasal also tend to be simultaneous in Onset position, but the velic opening precedes the oral closure in coda position, causing nasalization of the preceding vowel.

(10) [mĩm]



As Krakow (1999) points out, /l/ allophony has nothing to do with vowel nasalization in a featural view. But when represented gesturally, there is a clear parallelism: onsets are characterized by simultaneous production of gestures, and by strength of the oral gesture; codas are characterized by sequential production, with the oral gestures weaker and later. Findings like this have led to a strong focus on the role of syllable structure in gestural organization.

19.4 The coupling model of syllable structure

The recent AP focus on syllable structure has led to a new conception of the principles underlying gestural coordination. The older linguistic gestural model, where pairs of gestures were assigned relations like CENTER = ONSET (Gafos 2002), has been replaced by a theory in which there are only two kinds of gestural coupling (Saltzman & Byrd 2000; Nam & Saltzman 2003; Nam 2007; Nam et al. 2009).

As mentioned before, AP draws on a body of work on the coordination of skilled motion, such as limb oscillation (Turvey 1990). In this work, skilled motions are modelled as being similar to critically damped oscillators. An oscillator is a system that displays a periodic movement, like a pendulum, or a spring with a weight attached. “Critical damping” means that the oscillator slows down as it approaches the target. For a real-life analogy, think of the springs between a car’s chassis and frame: if you push on the bumper and then release it, the springs return it directly to its equilibrium position, but shock absorbers critically damp the springs so that the bumper won’t bounce. The dynamical equations that describe this kind of motion are similar to what the task dynamic model uses to describe speech motions. Once a task is activated (such as “tongue tip alveolar closure”), the articulator(s) start moving towards the target as if being controlled by a spring, but slow down on approach as if damped.

Each gesture is modelled as being controlled by a nonlinear planning oscillator, or “clock.” If we imagine a clock hand travelling through a 360° rotation, the beginning of the gesture occurs when the hand is at 0° phase and the end of the gesture at 360° phase.

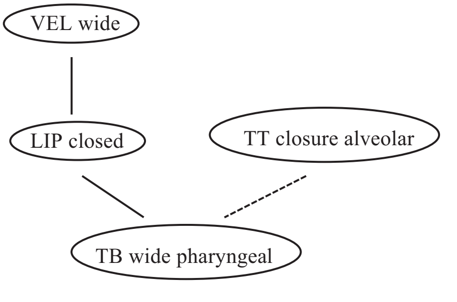
Oscillators can affect each other’s movement if they are “coupled.” This observation goes back to 1665, when Dutch physicist Christiaan Huygens noticed that the two pendulum clocks on his mantelpiece always beat in unison, and would return to this unison even if he deliberately disrupted their timing. He deduced that the clocks were subtly affecting each other through vibrations transmitted through the mantelpiece. Since then, physicists have shown that coupled oscillators tend to stabilize in one of two “normal modes”: in-phase timing, in which the oscillations are parallel (i.e., two pendulums swinging right and left at the same time), or anti-phase timing, in which the oscillations are opposite (i.e., one pendulum starting left when the other starts right). The process of gravitating towards these stable modes is called entrainment or mode locking.

Biophysicists have argued that entrainment is seen in the coordination of skilled motions as well (Haken et al. 1985). If you try to repeatedly do two movements at the same time (for example, tap your two index fingers on a table), you will tend to coordinate them either in-phase, by tapping the fingers simultaneously, or anti-phase, by alternating taps. Of these two modes, in-phase coordination is easier and more stable; the coordinated movements are very consistent in their relative timing and the rhythm is resistant to change. As a task gets harder, for example by speeding up the rate of finger-tapping, people tend to spontaneously switch from anti-phase to in-phase coordination.

Any "phase-lock" other than 0° or 180° relative phase is fairly difficult to maintain. People do accomplish more complex phasings when they learn skills like drumming or juggling, but these typically require considerable practice and often instruction.

It is hypothesized that speech evolved to use intrinsically stable modes of coordination whenever possible (Goldstein et al. 2006). Recent work (Saltzman & Byrd 2000; Nam & Saltzman 2003; Nam 2007; Nam et al. 2009) has pursued the hypothesis that *all* gestural coordination can be captured with just two phasing relations. If two gestures have a controlled timing relation, then they are coupled either in-phase or anti-phase. The input to the gestural model consists of a "coupling graph" specifying which gestures are coupled and how. For example, the graph below shows the coupling structure that Goldstein et al. (2007) hypothesize for English [mæd]. Solid lines indicate in-phase coupling (meaning that the gestures would ideally begin simultaneously); the dotted line indicates anti-phase coupling. This is not the only conceivable coupling graph for this word, of course, and whether it is the correct one is a question to be settled empirically.

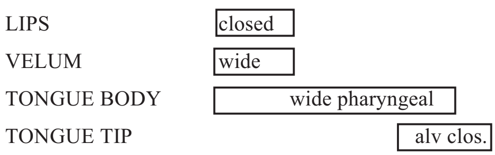
(11) Coupling graph of English *mad*, after Goldstein et al. (2007)



Once a coupling graph is established, it must be converted to a gestural score. This is done through an "intergestural level," which is implemented in TaDA. As described in Nam (2007), this level consists of a planning process that determines the most stable coordination of the entire gestural constellation. This is accomplished through a planning simulation in which each oscillator is started at an arbitrary phase of its clock (for example, VEL might begin at 40°, LIP at 65°, etc.), and all the gestures are set to oscillate repeatedly. At first, their relative timing changes on each repetition, as they are gradually pulled away from their random initial phasing relations towards the in-phase or anti-phase relations designated in the coupling graph. But eventually they settle into a stable pattern of relative timing, which stops changing from one repetition to the next. This stable timing pattern, plus a speech rate parameter, is the basis of the gestural score. (Again, this planning process takes the place of the earlier approach in which the gestural score was determined by explicit rules like CENTER = ONSET).

In the case of *mad*, the gestural score will be something like that below. This is a relatively simple case, because none of the in-phase or anti-phase coupling are in competition with one another (a problem that will be discussed further in section 19.4.1).

(12) Gestural score of English *mad*



These planning simulations have more than one role in theory: they identify the most stable timing pattern and output a gestural score, but they can also be used to compare the stability of different coupling arrangements. This is tested through adding a "noise factor" such as random variation in speech rate (Nam & Saltzman 2003: 2254), and seeing how much this disrupts the gestures' relative timing. Simulations can also yield a measure of stabilization time, or how long it takes for the stable pattern to emerge. For example, Nam (2007: 497) carried out simulations in which a consonant was considered to have separate gestures for its closure and its release. He found a faster stabilization time for the consonant's closure-release phasing in CV syllables than in VC. Faster stabilization time is assumed to correlate with faster planning time, from the speaker's point of view, and it is assumed that sequences with faster planning time will be preferred because they are easier to produce.

19.4.1 Coupling and syllable structure

There are several reasons to think that in-phase and anti-phase coupling may define the difference between syllable onsets and syllable codas.

First, this fits with the results of many articulatory studies (Browman & Goldstein 1988; Honorof & Browman 1995; Marin 2013; Pastätter & Pouplier 2014). It has been observed that gestures in a syllable Onset tend to start about simultaneously, both with one another and with the vowel gesture. Short lags of up to 50 ms. or so are common, but the numbers trend toward zero, as would be expected if they are in-phase. This is seen above in (12), where the TB gesture of the vowel begins around the same time as the VEL wide gesture and the LIP closure gesture. Coda gestures, on the other hand, start partway through the vowel gesture, consistent with an anti-phase relation, and if there are multiple gestures in the coda they tend to spread out rather than be produced simultaneously.

Second, onsets in real speech tend to show less variability in their timing than codas (Byrd 1996). This fits with the finding that in-phase timing is typically more stable than anti-phase (Haken et al. 1985; Goldstein et al. 2006). Nam & Saltzman (2003) show through simulations that adding a noise factor causes greater variability in codas than in onsets.

One intriguing implication of this approach is that it offers a new possible explanation of the well-known typological generalization that onsets are cross-linguistically preferred over codas. If onsets reflect in-phase coordination and in-phase coordination is easier (as studies outside linguistics propose), then it is not surprising that all languages allow onsets while many ban codas. It may also help explain why codas are typically acquired later by children; why the inventory and frequency of codas is typically lower than that of onsets;

why onset–nucleus combinations are very free while nucleus–coda combinations are often constrained; and why VC#V sequences are frequently resyllabified to V.CV. As Nam (2007: 489) observes, these patterns can all be captured with the generalization that languages prefer to maximize synchronous (in-phase) coupling, while minimizing asynchronous (anti-phase) coupling.

It should be noted that this theory does not explain all aspects of syllable typology: for example, it does not explain why languages disfavour onsetless syllables. Nor is this the only functional advantage proposed for CV syllables. Ohala (1996), for example, makes the case that onsets are easier to hear. These explanations are not mutually exclusive, of course; they may be mutually reinforcing.

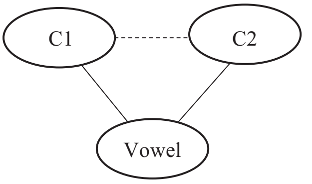
19.4.2 Complex onsets: the c-centre effect

The examples of coupling shown above were relatively simple, in the sense that none of the couplings were in competition with one another. But a more complicated situation arises when gestures have mutually conflicting coupling relations, as happens with certain combinations of multiple Onset or multiple coda gestures.

As noted above, Onset gestures are hypothesized to have an in-phase relation to the vowel. Having multiple gestures in-phase with the vowel is not a problem as long as those gestures can be produced simultaneously while still being perceptually recoverable. This is generally the case with the gestures that make up what is traditionally considered a segment, such as the tongue body and tongue tip gestures of an /l/, or the glottal opening and lip closure of a /p/.

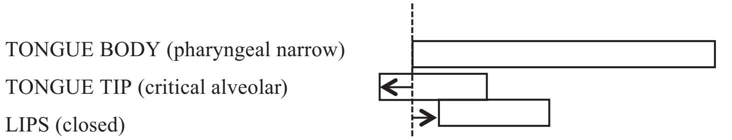
But in other cases, two Onset gestures would not be recoverable if they were produced simultaneously. For example, if an Onset contains a tongue tip critical gesture and a lip closure gesture (as in *spa*), producing them simultaneously would cause the tongue tip gesture to be acoustically masked by the lip closure. For both gestures to be recoverable, they must be in anti-phase relation to one another. This is shown below: the two consonantal gestures are coupled anti-phase with one another, but both are coupled in-phase with the vowel, since both are in Onset position.

(13) Coupling relations in CCV onset



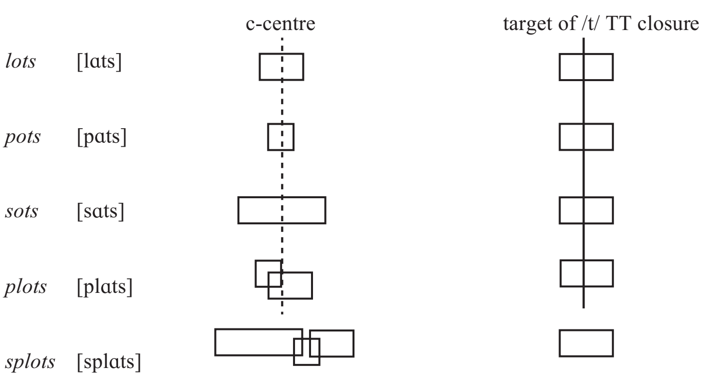
This coupling graph presents a problem: it is not possible for all three coupling relationships to achieve their target phasing. If both Onset gestures are perfectly in-phase with the vowel, they cannot be anti-phase with one another. However, Nam & Saltzman (2003) show that the planning simulation still does arrive at a stable timing pattern, which is a compromise between the desired phasings. As shown below, the first consonant begins before the vowel, and the second consonant begins after the vowel. Neither Onset gesture has exactly its preferred timing with respect to the vowel, as represented by the dotted line; each is shifted by about an equal distance, as represented by the arrows.

(14) Predicted timing of complex onsets; competitive coupled oscillator model



This is in fact what happens in real speech as well, at least in some languages. A series of studies of English (Browman & Goldstein 1988; Honorof & Browman 1995; Byrd 1995, among others) have shown that there is a stable relationship between the centre of the entire Onset (whether it consists of one, two, or three consonants) and the rest of the syllable. This is known as the c-centre effect. For example, Browman & Goldstein (1988) compared X-ray microbeam records of an English speaker articulating words like *lots*, *pots*, *plots*, and *splots*. They measured the difference between each Onset gesture and the target achievement of coda tongue tip raising of /t/. The coda gesture was chosen as an anchor because it was easier to identify than the target achievement of the vowel. The results are shown schematically below. With singleton onsets, there was a relatively consistent distance between the centre of the oral gestures and the coda; with complex onsets, the temporal midpoint of the whole Onset fell at around the same point in time that the midpoint of a singleton Onset would occupy. This is known as the c-centre effect, where "c-centre" (shown as a dotted line below) refers to the collective midpoint of the Onset oral gestures. Computing the distance between other anchors, such as the right or left edge of the Onset to the coda, yielded higher standard deviations.

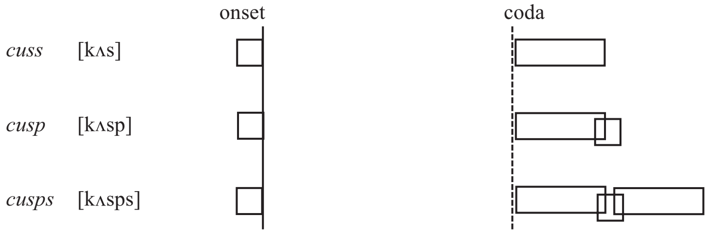
(15) Timing of oral gestures in onsets (Browman & Goldstein 1988)



The c-centre effect held only for consonants that formed an Onset. In a phrase like *piece plots* [pis plɑts], the tongue tip critical gestures of the first /s/ did not participate in the c-centre effect with respect to the following syllable.

Codas do not appear to participate in c-centre effects. Several studies of English syllables (Browman & Goldstein 1988; Honorof & Browman 1995) find that in codas there is a stable relationship between the left edge of the first coda consonant (shown as a dotted line below) and the rest of the syllable. As more consonants are added to the coda, the syllable simply becomes longer; the first coda consonant does not change its timing relative to the Onset and vowel. This suggests that coda consonants are coupled anti-phase with one another, and that only the first coda consonant is coupled with the vowel.

(16) Timing of coda clusters (based on Honorof & Browman 1995)



However, Byrd's (1995) study of five English speakers found some individual variation in the global timing of syllables, suggesting that not all speakers of a dialect necessarily use the same gestural organization.

It is an open question how many languages show this asymmetry in the timing of onsets and codas. The c-centre effect has been found for onsets in French (Kühnert et al. 2006), Italian (Hermes et al. 2013), and Georgian (Goldstein et al. 2007), but not for Slovak (Pouplier & Beňuš 2011: 18). Kochetov (2006) finds an onset–coda timing asymmetry in Russian, but it is different than the pattern in English. In some languages there seems to be variation depending on the type of cluster involved: Marin (2013) finds the c-centre effect in Romanian for sibilant-initial Onset clusters, but not stop-sibilant Onset clusters. Pastätter & Pouplier (2014) find similar patterns for sibilant-initial and sibilant-final Onset clusters in Polish. They suggest that sibilants may be resistant to overlap with the vowel, and that this could disrupt the c-centre effect for specific clusters. Both Romanian and Polish codas show sequential organization similar to that of English codas.

Of course, some cross-linguistic variability in the organization of gestures is not surprising, given that phonologists have long argued that languages differ in how they syllabify similar strings of sounds. Given the relatively small number of languages examined to date and the centrality of this topic for understanding gestural timing, this is likely to remain a central area of research in AP for the near future.

19.4.3 Physical study of syllable structure

Phonologists do not always agree on the syllabification rules of particular languages. One intriguing implication of the AP approach is that disputes about syllable structure could be settled empirically, through articulatory data. If we hypothesize that the traditional notion of "complex onset" refers to consonantal gestures that participate in a c-centre effect, then this structural pattern can be detected experimentally. Under this view, cross-linguistic differences in the timing of CCV sequences, discussed in the previous section, would be equivalent to cross-linguistic differences in the syllabification of such sequences. Of course, it remains to be shown that the c-centre effect consistently correlates with traditional, distributional diagnostics of complex onsethood. Yet several early results support the idea that it may.

In Moroccan Arabic, for example, there is controversy over the syllabification of consonant clusters in words such as /kra/ 'rent'. While some phonologists assume that /kr/ is a complex onset, there is evidence (especially from oral poetic meter) that only the /r/ is an ONSET consonant, and the /k/ has some other status. Proposals vary as to whether it is a "minor syllable," a syllable nucleus, or is licensed by a mora (see Shaw et al. 2009 for background). Shaw et al. (2009) studied the articulation of such sequences using EMA, comparing the results to simulations using TaDA. They were particularly interested in patterns of temporal stability, which is generally strongest within syllables. The timing patterns found for prevocalic CC clusters were most consistent with the hypothesis that C1 was a syllable nucleus, rather than a complex onset.

Goldstein et al. (2007) compared CV, CCV, and CCCV sequences in Georgian and Tashlhiyt Berber, using EMA. They found Georgian shows the c-centre effect, while Berber does not. In Berber, words like /mun/, /s-mun/, and /t-s-mun/ ('accompany', 'cause-accompany', '3FS-cause-accompany') all had the same relative timing of /m/ to /u/. Georgian is traditionally analyzed as having complex onsets, while Berber is usually analyzed as having only single-C onsets, so the phonetic findings accord with other evidence that these languages organize sounds differently.

These techniques can also be used to compare the organization of different gestural clusters within one language, as in a recent study of "impure s" in Italian. There are various arguments, both distributional and psycholinguistic, that Italian word-initial /sC/ clusters are different than other Italian CC clusters. For example, they condition a special allomorph of the definite article: *il sale*, *il premio*, but *lo studente*. Using EMA, Hermes et al. (2013) show that the c-centre effect holds for initial clusters such as /pr/, but not for /sp/. In /prima/, the /r/ shifts rightward compared to /rima/, but in /spina/, the /p/ has the same timing as in the name /pina/. This finding fits with other evidence that /s/ is not part of the syllable onset.

19.4.4 Moraic structure

The coupling model may shed light on another long-standing puzzle about syllable structure: why do coda consonants contribute to syllable weight in some languages but not others, and why do ONSET consonants never contribute to weight? One possibility is that the phonological patterning associated with "moraic" codas relates to a kind of timing relation. Nam (2007) attempts to model the difference between moraic and non-moraic codas. In his approach, every oral constriction involves a coupling of two gestures: a closure gesture and a release gesture. Nam hypothesizes that in languages where coda consonants add weight to the syllable, the vowel is coupled only with the closure gesture of the coda. In languages where coda consonants do not add weight, the vowel is coupled with both the closure and release gestures. The multiple couplings increase overlap between the vowel and coda, resulting in a shorter syllable. This proposal has yet to be rigorously tested against a range of languages, but offers an interesting hypothesis as to how gestural timing could relate to traditional notions of syllable weight.

19.5 New directions for the coupling model

19.5.1 Modelling phonological acquisition

Although the main application of the coupling model has been to understanding patterns of gestural organization cross-linguistically, several researchers have argued that the model also makes predictions about the acquisition and processing of phonological structure. As noted above, the model proposes that gestural scores are produced through a planning simulation, in which gestures oscillate repeatedly until they settle into a stable pattern of coordination ("entrainment"). Different gestural coupling structures require different numbers of oscillations to reach this stable phasing. It has been suggested that the time required for entrainment of a particular structure is a prediction both of how difficult speakers find it to acquire the structure, and how long it takes them to plan the production of the structure.

Nam et al. (2009) simulates acquisition of syllable structure in a Hebbian learning model in which a child agent tunes its initially random phase representations to match the perceived relative phase in adult productions. The adult's productions were varied across languages (for example, there are more tokens of codas in some simulations than others), to simulate the environment of languages with different frequencies of particular syllable structures. It was found that the child's CV phasing always stabilized faster than VC phasing. The lag is greatest in simulations where the adult produces more CV tokens, but strikingly, it persists even if the adult produces more VC tokens than CV tokens. This suggests that the greater ease of learning in-phase coordination can overcome even a paucity of such tokens in the environment. However, when VCC and CCV structures are added to the simulation, after acquisition of CV and VC, the child agent is quicker to master VCC than CCV. This is counterintuitive based on the idea that codas are "marked," but it follows from the fact that (in the simulation) VCC has a simpler phasing structure than CCV. CCV involves two in-phase couplings that are in competition; VCC involves two non-competitive anti-phase couplings. The simulation accords with reports that children have been found to acquire complex codas before complex onsets in some languages (see references in Nam et al. 2009: 2).

AP may also help explain why certain CV combinations are favoured in acquisition, and more frequent in the adult lexicon (Goldstein et al. 2006; Giulivi et al. 2011). During the babbling stage, children tend to produce CV syllables where the overlapping gestures are mechanically independent, like the lip and tongue body gestures in /ba/, or involve constrictions in similar locations, such as the two tongue body gestures in /gu/. Giulivi et al. (2011) use TaDA simulations to identify the most "synergistic" CV combinations, where synergy means that the final tongue body configuration for the C and V are similar. They argue that this measure of synergy predicts how easy it will be to produce the C and V in-phase.

19.5.2 Gestural coordination and morphological structure

The coupling model may offer a new approach to the phonology of morpheme boundaries. There are some indications that gestures belonging to a single lexical entry are coordinated in a different way than gestures that belong to different morphemes.

Cho (2001) used EPG to study morphologically simplex and compound words in Korean. He found that a sequence such as [ti] showed more variability in the relative timing of the oral gestures associated with /t/ and /i/ when it was heteromorphemic, as in /mat-i/ 'the oldest', than when it was mono-morphemic, as in /mati/ 'knot'. A similar difference is found between lexicalized and non-lexicalized compounds. Cho proposes that this is because the timing relations are lexically specified in 'knot', where the /t/ and /i/ gestures are part of single lexical entry. In Cho's Optimality Theoretic (OT) analysis, the lexically specified timing relation is stronger because it is protected by IDENT constraints.

Nam & Saltzman (2003), in a non-OT analysis, demonstrate through simulations that this difference variability could follow from the different coupling patterns below. When an intervocalic C is not with the following V, the two will show more inconsistent timing than if they had a specified coupling relation.

(17) Coupling within and across morphemes

/ati/ (same morpheme) /at-i/ (different morphemes)

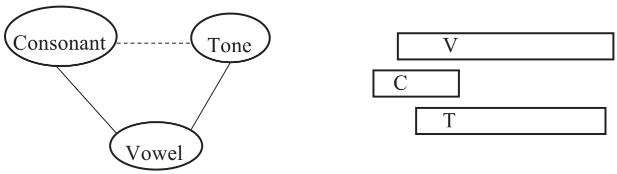
It is possible that even some categorical morphophonological alternations could be reanalyzed as effects of coordination. Goldstein (2011) reanalyzes English past tense allomorphy, a classic case of apparent segmental alternations, in terms of gestural coordination. He suggests that the three reported past tense allomorphs (-t, -d, -ɨd) actually consist of the same gestures, namely a TT closure and release, plus a VEL closure to prevent nasality. Through simulations, he shows that a natural-sounding output can be achieved for words like *nabbed* and *napped* [næbd, næpt] by coupling the TT release gesture of the suffix to the release gesture of the preceding consonant. In *napped*, the glottal opening gesture associated with the /p/ inhibits voicing on the suffix (to a lesser extent, the same happens with *nabbed* simply due to the length of the closure). Under this proposal, there is no phonological alternation and no allomorphy in such words: whether the suffix sounds like [-t] or [-d], it consists of the same gestures in the same coupling relations. As for the [-ɨd] variant, Goldstein proposes that the suffix still consists of the same gestures (with no vowel gesture for the [ɨ]), but they participate in a different coordination relation with less overlap between Cs. This creates the percept of a transitional targetless vocoid (as discussed relative to schwa in section 19.3.2). EMA and real-time MRI studies confirm that tongue body shapes during the [-ɨd] suffix are consistent with lack of TB target (Smorodinsky 2001; Lammert et al. 2014). Goldstein's simulations show that if the tongue tip closure of the suffix is coupled with stem-final closure gesture, no vocoid occurs; but if the tongue tip closure of the suffix is coupled to the stem-final release gesture, a vocoid does occur.

19.5.3 Coupling models of tone and intonation

Recently, the coupling model has been extended to account for tone and intonation. This area was pioneered by Gao's (2009) work on Mandarin tone, which argues that tonal movements can be described as gestures, and that tonal gestures engage in phasing relations with one another as well as with vocalic and consonantal gestures. Just as a typical C or V gesture is a task of reaching a constriction target, a tonal gesture is a task of reaching a tonal target. Unlike other constriction gestures, the presence of a tonal gesture can be read directly from the acoustic record. A High tone gesture, for example, begins at the ONSET of a pitch rise, and ends when the highest point is achieved. In giving tones a duration, this approach differs from autosegmental-metrical theory (Goldsmith 1990), in which tones are thought of as dimensionless points, with intervening time periods filled in by interpolation.

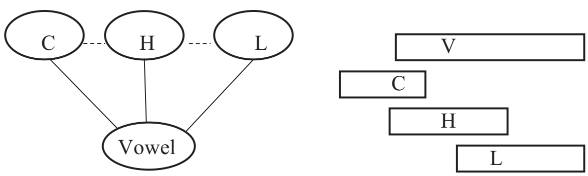
Gao (2009) (as described in Mücke et al. 2012) proposes that in a CV syllable with a single lexical Tone (T), such as Mandarin Tone 1 (High) and Tone 3 (Low), the T gesture behaves essentially like an additional ONSET C gesture. The C and T gestures are coordinated anti-phase with one another, but both in-phase with V gesture. This causes a shift in alignment exactly analogous to the c-centre effect, so that the gestures are actually activated in the order C, V, T. The C and T gestures each begin about 50 ms from the V gesture (whose ONSET is identified from articulatory records as usual).

(18) Coupling relations in lexical tone on a CV syllable, after Gao (2009)



For complex tones, like Mandarin Tone 4 (High-Low), the same principle applies. The gestures of C, H, and L are coupled anti-phase to one another, but in-phase to the vowel. The H gesture begins simultaneously with the V gesture, while C is pushed earlier and V later.

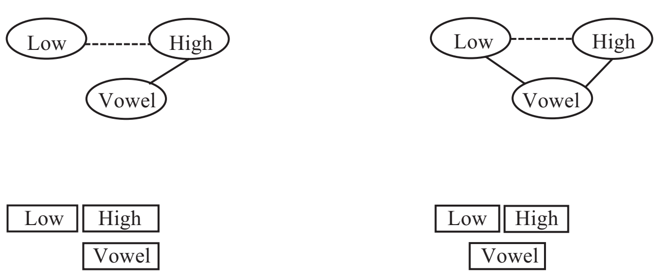
(19) Coupling relations in complex lexical tone



Gestural phasing can also be used to analyze the coordination of intonational tones. Mücke et al. (2012) offer a gestural analysis of intonational rises in Catalan- and German-stressed syllables. They found that in Catalan, the C, V, and T gestures start about simultaneously – there was no c-centre effect like that Gao (2009) found in Mandarin. In German, C and V begin together but the rise starts much later. Modelling with TaDA shows that the difference between German and Catalan can be captured by different couplings: in both languages, L and H tones are anti-phase with one another and H is in-phase with V; in German, L is also in-phase with V, while in Catalan there is no coupling of L and V. The graphs below show the gestural scores predicted by each coupling relation.

(20) Coupling relations and gestural scores in intonational tone, after Mücke et al. (2012)

(a) Catalan (b) German



Cross-linguistic differences in the realization of intonational rises may also result from different tonal compositions. Niemann et al. (2011) argue that while rises in German reflect two gestures, Low and High, as shown above, rises in Italian reflect only a single high gesture.

Mücke et al. (2012) propose an interesting hypothesis: they suggest that effects like that in Mandarin, where a tone participates in the c-centre effect, only likely occur in lexical tone systems, where coupling between tone and non-tone gestures is represented lexically. Non-lexical tones, like those of German and Catalan, are unlikely to affect within-syllable coupling relations.

19.6 Summary

Over the past thirty years, the AP approach has been applied to an increasingly wide range of problems in sound structure. Although the number of languages studied still remains small, and many topics such as morphology and intonation are only beginning to receive attention, recent work in these areas show promise for the development of a more comprehensive model of speech, including cross-linguistic variation. AP research holds itself to unusually rigorous empirical standards, generally demanding that analyses be based on precise articulatory records and computationally explicit simulations of speech production. It is unique among phonological frameworks in the extent to which it draws on, and participates in, a wider tradition of work on biomechanics. A biophysicist wandering into a linguistics conference would not recognize the abstract entities posited in most phonological frameworks (moras, archiphonemes, faithfulness constraints, etc.), but s/he would understand what it means to model speech movements as coupled oscillators.

19.7 Further reading

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 341–376.
One of the earlier works on AP, this article introduces basic concepts and demonstrates how casual speech processes can be described in terms of changes in gestural magnitude or overlap between gestures.

Gafos, A. I. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, *20*(2), 269–337.
This article shows how AP representations can be used within an Optimality Theoretic (OT) grammar. The grammar sketched focusses on the coordination of CC sequences in Moroccan Arabic.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. *Approaches to Phonological Complexity*, 299–328.
This article typifies the more recent AP approach to syllable structure, arguing that CV syllables are unmarked because they result from in-phase coupling of C and V gestures. It provides a good introduction to the concept of coupled oscillators.

19.8 Related topics

Chapter 18: Laboratory Phonology

References

Bernstein, N. A. 1967. *The co-ordination and regulation of movements*. London: Pergamon.

Browman, C. P., & L. Goldstein. 1992a. 'Targetless' schwa: An articulatory analysis. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody* (pp. 26–56). Cambridge: Cambridge University Press.

Browman, C. P., & L. Goldstein. 1992b. Articulatory phonology: An overview. *Phonetica*, *49*(3–4), 155–180.

Browman, C. P., & L. Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.

Browman, C. P., & L. Goldstein. 1989. Articulatory gestures as phonological units. *Phonology*, *6*(2), 201–251.
Browman, C. P., & L. Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2–4), 140–155.
Browman, C. P., & L. M. Goldstein. 1986. Towards an articulatory phonology. *Phonology*, *3*(1), 219–252.
Byrd, D. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*(2), 209–244.
Byrd, D. 1995. C-centers revisited. *Phonetica*, *52*(4), 285–306.
Cho, T. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, *58*(3), 129–162.
Davidson, L., & M. Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. In *Proceedings of the West Coast Conference on Formal Linguistics* (Vol. 22, pp. 165–178).
Gafos, A. I. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory*, *20*(2), 269–337.
Gao, M. 2009. Gestural coordination among vowel, consonant and tone gestures in Mandarin Chinese. *Chinese Journal of Phonetics*, *2*, 43–50.
Gick, B., F. Campbell, S. Oh, & L. Tamburri-Watt. 2006. Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*, *34*(1), 49–72.
Gick, B., & I. Wilson. 2006. Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. *Laboratory Phonology*, *8*, 635–659.
Goldsmith, J. A. 1990. *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.
Goldstein, L. 2011. Back to the past tense in English. In R. Gutierrez-Bravo, L. Mikkelsen, & Eric Potsdam (Eds.), *Representing language: Essays in honor of Judith Aissen* (pp. 69–88). UCSC Linguistics Research Center.
Goldstein, L. 1994. Possible articulatory bases for the class of guttural consonants. In P. Keating (Ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonology III* (pp. 234–241). Cambridge: Cambridge University Press.
Goldstein, L., D. Byrd, & E. Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215–249). Cambridge: Cambridge University Press.
Goldstein, L., I. Chitoran, & E. Selkirk. 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. In *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp. 241–244).
Goldstein, L., M. Pouplier, I. Chen, E. Saltzman, & D. Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition*, *103*(3), 386–412.
Haken, H., J. S. Kelso, & H. Bunz. 1985. A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, *51*(5), 347–356.
Hall, N. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology*, *23*(3), 387–429.
Hanson, H. M., & K. N. Stevens. 2002. A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using HLsyn. *The Journal of the Acoustical Society of America*, *112*(3), 1158–1182.
Hawkins, S. 1992. An introduction to task dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody* (pp. 9–25). Cambridge: Cambridge University Press.
Hermes, A., D. Mücke, & M. Grice. 2013. Gestural coordination of Italian word-initial clusters: The case of 'impure s'. *Phonology*, *30*(1), 1–25.
Honorof, D. N., & C. P. Browman. 1995. The centre or edge: How are consonant clusters organized with respect to the vowel. In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Vol. 3, pp. 552–555).
Iskarous, K., J. McDonough, & D. H. Whalen. 2012. A gestural account of the velar fricative in Navajo. *Laboratory Phonology*, *3*(1), 195–210.
Kochetov, A. 2006. Syllable position effects and gestural organization: Evidence from Russian. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in Laboratory Phonology VIII* (pp. 565–588). Cambridge: Cambridge University Press.

Krakow, R. A. 1999. Physiological organization of syllables: A review. *Journal of Phonetics*, *27*(1), 23–54.
Kühnert, B., P. Hoole, & C. Mooshammer. 2006. Gestural overlap and C-centre in selected French consonant clusters. In *Proceedings of the 7th Speech Production Seminar*.
Lammert, A., L. Goldstein, V. Ramanarayanan, & S. Narayanan. 2014. Gestural control in the English past-tense suffix: An articulatory study using real-time MRI. *Phonetica*, *71*(4), 229–248.
Marin, S. 2013. The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics*, *41*(3), 211–227.
Mücke, D., H. Nam, A. Hermes, & L. M. Goldstein. 2012. Coupling of tone and constriction gestures in pitch accents. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, & B. Kühnert (Eds.), *Consonant clusters and structural complexity* (pp. 205–230). Berlin: Mouton de Gruyter.
Munhall, K. G., E. Vatikiotis-Bateson, & Y. Tohkura. 1995. X-ray film database for speech research. *Journal of the Acoustical Society of America*, *98*, 1222–1224.
Nam, H. 2007. Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. *Laboratory Phonology*, *9*, 483–506.
Nam, H., L. Goldstein, C. Browman, P. Rubin, M. Proctor, & E. Saltzman. 2006. *TaDA (Task Dynamics Application) manual*. Available at haskins.yale.edu.
Nam, H., L. Goldstein, & E. Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupé (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin: Mouton de Gruyter.
Nam, H., L. Goldstein, E. Saltzman, & D. Byrd. 2004. TaDA: An enhanced, portable Task Dynamics model in MATLAB. *The Journal of the Acoustical Society of America*, *115*(5), 2430–2430.
Nam, H., & E. Saltzman. 2003. A competitive, coupled oscillator model of syllable structure. In *Proceedings of the 15th International Congress of Phonetic Sciences* (Vol. 1, No. 2253–2256).
Niemann, H., D. Mücke, H. Nam, L. Goldstein, & M. Grice. 2011. Tones as gestures: The case of Italian and German. In *Proceedings of ICPhS*, *XVII* (pp. 1486–1489).
Ohala, J. J. 1996. Speech perception is hearing sounds, not tongues. *The Journal of the Acoustical Society of America*, *99*(3), 1718–1725.
Pastätter, M., & M. Pouplier. 2014. The temporal coordination of Polish onset and coda clusters containing sibilants. In *Proceedings of the 10th International Seminar on Speech Production, Cologne, Germany, 5–8 May 2014*.
Pouplier, M., & Š. Beňuš. 2011. On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology*, *2*(2), 243–273.
Pouplier, M., & L. Goldstein. 2005. Asymmetries in the perception of speech production errors. *Journal of Phonetics*, *33*(1), 47–75.
Proctor, M., R. Bundgaard-Nielsen, C. Best, L. Goldstein, C. Kroos, & M. Harvey. 2010. Articulatory modelling of coronal stop contrasts in Wubuy. In *Proceedings of the 13th Australasian International Conference on Speech Science and Technology* (pp. 90–93).
Ramanarayanan, V., L. Goldstein, & S. S. Narayanan. 2013. Spatio-temporal articulatory movement primitives during speech production: Extraction, interpretation, and validation. *The Journal of the Acoustical Society of America*, *134*(2), 1378–1394.
Saltzman, E., & D. Byrd. 2000. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499–526.
Saltzman, E. L., & K. G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.
Shaw, J., A. I. Gafos, P. Hoole, & C. Zeroual. 2009. Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation. *Phonology*, *26*, 187–215.
Smorodinsky, I. 2001. Schwas with and without active gestural control. *The Journal of the Acoustical Society of America*, *109*(5), 2446–2446.
Sproat, R., & O. Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, *21*(3), 291–311.
Turvey, M. T. 1990. Coordination. *American Psychologist*, *45*(8), 938.
Zsiga, E. C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. *Phonology and Phonetic Evidence: Papers in Laboratory Phonology*, *4*, 282.

20

Exemplar theories in phonology

Stefan A. Frisch

The earliest formal theories in generative phonology argued that the nature of phonological computations and representations set language apart from other domains of cognition (Chomsky & Halle 1968). Since then, however, there have been a variety of advances in the theories of cognitive psychology in representation and processing of categories. One of these theories involves the use of individual memory traces in the representation of categories and the processing of stimuli, including novel stimuli. This is known as exemplar theory (see Nosofsky 2014 and the references therein for a recent review and explicit model). Exemplar theory is a theory of the representation and processing of categories in which stimuli are processed by comparing them to a set of previous experiences stored in memory. The goal of this chapter is to provide an introduction to exemplar theory as it has been applied in phonology.

20.1 Background, principles, and assumptions of exemplar theory

20.1.1 Historical introduction: indexical effects and network structures

The use of exemplar concepts and theories in phonology emerged from theories and findings in two closely related components of linguistics, phonetics and morphology. In phonetics in the 1950s and 60s, advances in technology allowed for quantitative measures of phonetic dimensions of phonological categories such as vowel formant frequencies and voice Onset time (e.g. Peterson & Barney 1952; Klatt 1975). The study of phonetics at this time generally assumed that segments in different phonological categories would be distinct along some acoustic dimension or set of dimensions that can be connected to the mapping from articulation to acoustics (Stevens 1989). The mapping is not necessarily straightforward, however, as raw acoustic measures show considerable overlap across speakers and within speakers depending on context. For example, a replication of Peterson & Barney (1952) using modern analysis techniques found the same F1 and F2 values could be found for /æ ɛ ʊ ɝ/ for different speakers (Hillenbrand, Getty, Clark, & Wheeler 1995).

Early theories of speech perception proposed that indexical information, related to the identity of the talker, had to be removed or normalized in some way in order to uncover the cues to a phonetic category (Klatt 1979). However, as researchers sought to investigate the normalization process it was instead found that speaker-specific information contributed to (or detracted from) the ability to perceive speech and recognize words (Mullennix & Pisoni 1990). Further, it was shown that speaker-specific information and lexical information were shown to be stored together (Goldinger 1996) and that when the perceivers became producers, episodic details of the perceived words had an impact on the speaker's productions (Goldinger 2000). These episodic effects persist over time and so cannot be a short-term memory phenomenon. In Goldinger (1996) experiment 2, participants showed talker-specific influences in word recognition a week after exposure. This shows that details of language experiences are stored in long-term memory as they survive the processes of memory consolidation that happen during sleep (e.g. Rasch & Born 2013).

In roughly the same time period, work by functional linguists in morphology was exploring a network connection model of lexical structure for morphology (Bybee 1988), presumably inspired in part by cognitive psychology work in connectionism. This work was related to the general approach in functional linguistics to seriously consider notions like category similarity and analogy as predictors of language patterns (e.g. Skousen 1989).

While formal generative linguistics has symbolic representations and transformational rules defined by a logic structure, Bybee (1988) noted that morphologically related forms within a paradigm are not all created equal. For example, cross linguistically the present tense second and third person forms are more closely related to one another than to their past tense forms. They are more likely to have similar allomorphs and to affect one another in historical change. More generally, high frequency word forms are usually the unmarked or least marked forms. Further, the productivity of morphophonological patterns is related to how common they are across the lexicon. All of these various dimensions of within-category structure go beyond what can be represented in a symbolic description or coded in a formal SPE-style generative rule. Bybee (1988) therefore proposes that morphologically complex forms are stored in the lexicon with structural and conceptual connections to one another.

Taken together, lexical storage of both phonetic detail and morphologically complex words sets the stage for phonology to also show patterns compatible with an exemplar theory. Bybee (1988) already discussed within-category effects of the productivity of morphophonological rules. Bybee (2001) reviews a variety of phonological phenomena that would be unexplained in a generative theory because they require a domain of generalization that is different from an individual lexical item or a symbolic rule. One of these phenomena concerns the influence of the frequency of use of a lexical item on its phonetic or phonological structure. For example, English words that possibly contain a medial rhotic vowel show a range of patterns. Some words, like *every*, sound unusual if pronounced /ɛvɚi/ with a medial syllable rather than as bisyllabic /ɛvri/. For others, like *mammary*, it is the reverse. Bybee proposed that word usage has an impact on lexical representation such that frequently used words are reduced. This theory predicts that low frequency words will more often be produced with a rhotic vowel, while high frequency words will have a rhotic Onset. Bybee discusses a similar pattern with flapping of /t d/, and these findings have been replicated in larger studies and in other languages (Jurafsky, Bell, Gregory, & Raymond 2001; Pluymaekers, Ernestus, & Baayen 2005). These examples of frequency effects in the lexicon are related to usage frequency, also known as *token frequency* of the word.

In addition to influences of the token frequency of lexical items on their phonetic or phonological structure, in an exemplar model the frequencies of the phonological structures themselves can play a role in language sound structure. In principle, there is no reason to distinguish generalizations at the level of lexical items or segments from generalizations at other intermediate levels (or, for that matter, sub-segmental levels or phrasal levels). As one example, Bybee & Pardo (1981) found that a potential morphophonological rule for stem vowel changes in Spanish was highly represented among verbs with front vowel /e/ but rare with verbs with back vowel /o/. In a nonce verb task, many participants demonstrated use of stem vowel change in the front vowel, but only one used it with a back vowel. While a generative model could easily describe the pattern either as one affecting all mid vowels or only the front mid vowel, the status of the pattern for the back mid vowel is unexplained in either case. The rule is unable to describe the possibility but unlikelihood that a participant will generalize a model based on analogical extension of pattern in the lexicon, and so an exemplar model among others, can explain both the productivity in one case and the marginal productivity in another. In the theory, analogical extension of a pattern will only occur if it is present in a variety of words. Frequency in terms of the number of words containing a pattern is known as *type frequency*.

20.1.2 Principles of exemplar theory

As noted above, exemplar theory is a theory of the representation and processing of categories in which stimuli are processed by comparing them to a set of previous experiences stored in memory. Each instance stored in memory is an EXEMPLAR, and crucially the nature of what is stored and how it is compared to a novel stimulus is important. Exemplar theories can be differentiated from other theories of categorization in cognitive psychology in that the exemplars are taken to be experiences that include a relatively rich amount of detail. In the case of phonology, then, these rich experiences would include variant productions of vowels with their different formant frequencies due to differences in context and talker, among other factors. For lexical items, individual experiences of the lexical items are produced by different talkers, in different prosodic positions, at different speech rates, and potentially even within different listening conditions. While researchers in linguistics would generally prefer to limit the information considered relevant to linguistic dimensions, there is no reason to think this is so in an exemplar theory. An anecdote that usually resonates with readers is the idea of wanting to return to reread an important point in a paper. In addition to searching for the desired text on the basis of key words (or pieces of linguistic information), the reader usually also has some idea of where on the page the text was or its placement relative to a section heading or graphic. That type of contextual detail is completely unrelated to the linguistic information, but nonetheless the mechanisms of the perceptual system register that information in some way. Over the course of multiple experiences, meaningful covariation will become a part of the knowledge of that category, and meaningless covariation will be noise that does not.

The processing of categories in an exemplar theory involves comparing a new stimulus with previously stored exemplars. If the new stimulus is similar to a number of exemplars from a particular category, then the stimulus is likely to be considered a member of that category. The notion of similarity plays a crucial role in the comparison process, and so the appeal of an exemplar theory to linguists exploring the role of analogy in linguistic knowledge is straightforward. Similarity itself is a complex notion, and there are diverse issues in similarity that are studied in cognitive psychology. It is generally believed that similarity is determined by proximity in a multidimensional space of features or concepts (Gärdenfors 2000; Nosofsky 2014). This can work quite well for phonology where the study of distinctive features and the relations between features and phonetic characteristics has been going on for quite some time. The features provide a ready basis for a phonological similarity space for segments (Frisch, Pierrehumbert, & Broe 2004). There are many ways a segmental similarity measure could be scaled up to whole word computations, but the problem has not been investigated in depth. Coleman & Pierrehumbert (1997) provide a model for computing overall word probability, which is one measure of similarity to the lexicon, taking into account onsets, rimes, position relative to word edges, and stress. Bailey & Hahn (2005) investigate similarity judgments between words and find evidence for computation involving phonological features of segments that is not the same as perceptual confusability. These two approaches can conceptually be united using a general psychological model of feature matches in structured representations such as in Goldstone (1994), but that work has not, to date, been done.

In addition to detailed representations and mapping within a similarity space, an exemplar approach to phonology includes the notion of category typicality as part of the theory. With categories built from groups of related exemplars, a newly encountered stimulus may be similar to relatively many or few exemplars in the category. As a result, an exemplar theory can account for gradience in wordlikeness judgments that are difficult to explain in a generative grammar (see, for example, papers in Fanselow, Féry, Schlesewsky, & Vogel 2006). In the case of phonological categories, the notion of typicality might be used to explain asymmetries in representation or processing related to the linguistic concept of markedness (Pierrehumbert 2016). For phonological or morphophonological patterns like the stem vowel change in Bybee & Pardo (1981), typicality relates to the number of items displaying the pattern.

In general studies of exemplar theories, exemplar models are also influenced by frequency of exposure and recency of exposure. Frequent or recent exposure to an exemplar member of a category makes that exemplar more easily accessible. A novel stimulus that is similar to that exemplar would then more readily activate the category. While it is less apparent how recency might apply in a synchronic phonological analysis of language patterns, there is some evidence to support this aspect of exemplar theory as well. Speakers clearly have this capacity, as seen in the ability to use a novel language pattern in language play, such as in the Phteven and Ermagerd girl memes. Language change as a result of language contact involves replacing long-established phonological patterns with recently acquired ones. For example, Carvalho (2006) argues that /s/-aspiration is being adopted as a prestige form by some social groups living along the border between Uruguay and Brazil. Phonological language change has been documented within an individual speaker showing that recency effects do influence phonological categories (Harrington, Palethorpe, & Watson 2000). A related but far less interesting finding comes from Frisch, Large, & Pisoni (2001). In a nonword judgment task, they note that in pilot experiments, participants gave slightly higher wordlikeness judgments to nonwords when the stimuli were played to listeners twice rather than once.

20.2 Issues and directions of exemplar theory

20.2.1 Probabilistic phonotactics

One of the strongest pieces of evidence for lexical or exemplar theories of phonological knowledge comes from studies of phototactics that show that the statistical distribution of segmental or suprasegmental categories across the lexicon can be systematic and are known to native speakers. In other words, one component of the knowledge of language sound structure is type frequency. In one of the first studies, Pierrehumbert (1994) examined the occurrence of medial three consonant clusters in English words. Pierrehumbert considered the generative hypothesis that attested medial clusters should be derivable from the possible Onset and coda clusters. She notes that with proper phonological treatment of the combination of a syllable coda and a following syllable onset, there are 8,708 possible medial triconsonantal clusters in English. However, a dictionary search finds only 50 attested clusters. By extending the information about onsets and codas to include their frequency of occurrence in the dictionary, the vast majority of the missing clusters can be accounted for. The attested clusters are for the most part among the 200 most likely combinations of codas with onsets based on type frequency. In other words, the triconsonantal clusters that are found in a dictionary search are the ones that combine high frequency constituents. In retrospect, this seems intuitively obvious. A novel word–generating algorithm that is stochastic, using the probabilities of different constituents based on their frequency of occurrence, would generate a space of novel nonwords similar to the existing lexicon.

The psychological reality of a stochastic phonotactic grammar was demonstrated in experimental studies with participants making judgments of wordlikeness or acceptability for novel nonwords (Coleman & Pierrehumbert 1997; Frisch, Large, & Pisoni 2000). The study by Coleman & Pierrehumbert (1997) was notable in that some of the Onset constituents used in the nonwords were unattested consonant clusters (such as Onset /mr/). While well-formedness judgments for novel nonwords including unattested clusters were lower than those containing only attested clusters, there were differences between nonwords containing unattested clusters that were predictable from the aggregate constituent frequencies of the remaining constituents. Frisch, Large, & Pisoni (2000) used novel nonwords consisting only of attested onsets and rimes and replicated the predictive value of statistical patterns in the lexicon on both wordlikeness and acceptability judgments. This finding was also extended to novel Spanish nonwords with the same stochastic grammar (Frisch & Brea-Spahn 2010).

In many cases, however, the statistical phonotactic results could alternatively be interpreted on the basis of similarity to existing words (Bailey & Hahn 2001). It is possible that both of these influences exist with similarity to existing words more relevant where many of these similarities exist, as in the dense regions of the phonotactic space. In sparse regions of the phonotactic space where there are few or no similar neighbors, phonotactic probability is more relevant (Frisch, Large, Zawaydeh, & Pisoni 2001).

Critics of these statistical models of phonology point to data where speakers appear to differentiate their acceptability of patterns that are non-existent, and therefore not represented by any exemplar or have a zero frequency. For example, Moreton (2002) compares the relatively acceptability of novel words with unattested [bw] and [dl] clusters. Indeed, these clusters have zero frequency in English word onsets. In experiments with novel nonwords, native English speakers show a preference for [bw] onsets over [dl] onsets. Moreton eventually concludes that feature level generalizations are required in order to properly capture differences between these clusters as a constraint on place. While at first glance this criticism might appear valid, it is overly simplistic. There is nothing inherent about statistical generalizations over the lexicon that requires the generalizations to apply to the Onset cluster level of representation only. A theory in which generalizations are emergent over patterns in the lexicon, like an exemplar theory, could discover those generalizations at a variety of levels of granularity (Goldrick 2002). A recent replication of these experimental findings and expanded analysis provide support for generalization from experience with phonological features over a novel lexicon (Cristia, Mielke, Daland, & Peperkamp 2013).

In a related study, Daland et al. (2011) collected judgment data on novel nonwords with a variety of initial consonant clusters that spanned a range of lexical and phonological conditions. Some Onset clusters were commonly attested (e.g. /tr/), others were marginally attested but rare or found only in obvious nonwords (e.g. /ʃl/), and a final set were unattested but varied in sonority contour according to cross-linguistic patterns (e.g. good sonority Onset /zr/ and poor sonority Onset /rg/). Participant ratings of novel nonwords with these clusters matched expected patterns, with attested clusters preferred over marginal clusters, preferred over unattested clusters. Within unattested clusters, sonority sequencing generalizations were reflected in participant judgments. Daland et al. considered a variety of phonotactic models based on generalization over the lexicon as predictors of participant judgments. Models included the stochastic model of Coleman & Pierrehumbert (1997) and simpler statistical models as well as two varieties of models designed to learn based on phonological features. The modeling results clearly showed that syllabic structure is required to model participant judgments. However, no model was able to accurately predict variation in rating for both attested clusters and novel clusters. The probabilistic models were more effective with attested clusters, while the phonological learning models with feature representations were more effective for novel clusters. Deland et al. conclude that sonority effects in unattested consonant clusters can be modeled with generalization over the lexicon if the models contained enough information to represent sonority generalizations, such as phonological features and syllabic structure.

20.2.2 Frequency in morphology and productivity

There is a variety of evidence in support of an exemplar-type theory of morphological structure. Studies of irregular verb morphology such as *ring*, *rang*, *rung* have argued that verbs can be grouped along dimensions of phonological similarity into schema or analogical groups that share patterns (e.g. Wang & Derwing 1994). The set of items included in a group here is a degree of generalization that is in between the usual dichotomy of representation in generative phonology, where forms either pattern according to a rule or are listed individually as lexical exceptions. These intermediate generalizations are not unique to English morphophonology. Analogous sets of sub-patterns have been found, for example, in phonological conditioning for verbal morphology in Spanish (Albright 2003) and in nasal substitution in Tagalog (Zuraw 2008) discussed in the next section. In this case, what might appear in a traditional analysis as arbitrary, and therefore lexically specified on an individual basis, can be seen as systematic when the structure of the lexical groups is less rigidly defined. An important difference between arbitrary lexical specification and the schema analysis is that the schema analysis makes predictions about the behavior of participants when processing novel items such as forms borrowed into the language, adapted ad hoc in code switching, or for nonce verbs used in an experimental setting. In addition, the schema or analogy approach is another case where the theoretical perspective can be applied at a variety of levels of linguistic generalization, not just for phonology. For example, Bybee & Eddington (2006) show that semantically related Spanish verb classes can be grouped into similarity-based analogical classes. They further argue that these classes have a prototype structure with a set of distinct central members that anchor the different classes with other forms related to the central members by similarity. In this case, then, the structure of the linguistic category is no different from the structure of a variety of other cognitive categories outside of language (Rosch 2003).

There is also evidence that indicates that morphological productivity is an emergent property from generalization over the lexicon, including consideration of type and token frequencies of morphemes (Hay & Baayen 2002; Hay 2003; Hay & Baayen 2003). However, the relationship between type and token frequency and morphological productivity is complex.

In order for a morpheme to be productive, it must be decomposed by the language learner. If it is not clear that a morpheme makes a sub-part of a larger unit, then the larger unit will be stored as a whole lexical item (for example, but it would seem the same ideas would apply to idiomatic expressions, and indeed there has been an argument for an exemplar model of patterns within some idiomatic constructions: Gries 2011).

Given the relationship between phonological generalization and type frequency, it would be expected that productivity would be related to the number of morphologically complex forms containing an affix that are in the lexicon. Once again, however, a morphologically complex form has to be decomposed in order for the affix to be apparent. Researchers in cognitive psychology have proposed that token frequency plays a role in decomposition such that high frequency morphologically complex words would be stored as wholes (Stemberger & MacWhinney 1986). In an extensive study of morphology in the English lexicon, Hay (2003) found that decomposition of morphologically complex forms is predictable from the token frequency of the stem as a monomorphemic word relative to the token frequency of the morphologically complex word. If the stem is higher in token frequency than the derived form, the derived form will be decomposed. For example, a form like *illegible* would not favor decomposition as it is more frequent that the stem form *legible*. In a purely symbolic model of morphology this notion is difficult to represent; in an exemplar model of morphology the similarity between morphemes and morphologically complex words can be gradient. Even in the network analysis of Bybee (1988), where morphemes are related by associative links, it is possible to have links of different strength of association.

The parsing analysis of Hay (2003) was revised slightly in Hay & Baayen (2002) by comparing the effect of stem and whole item frequency on parsing to psycholinguistic data. Using data from a lexical decision experiment in Dutch that included morphologically complex words, they concluded that there is a whole word bias in parsing such that morphologically complex forms are preferentially treated as whole words. This is compatible with findings in language acquisition unrelated to morphology (e.g. Dahan & Brent 1999). This adjustment in the estimate of affix parsing is combined with the notion that complex items can be decomposed to different degrees in parsing to refine the estimate of the number of morphologically complex words in the lexicon that are decomposed for each affix. Hay & Baayen (2002) demonstrate that the productivity for an affix depends on the number of items in the lexicon where it is likely to be parsed – in other words, the type frequency of the morphological combination.

Hay & Baayen (2003) further extend this analysis to consider the contribution of phonotactic probabilities to morphological parsing. Lexical parsing from a speech stream has been shown to consider phonotactic probabilities (e.g. Saffran, Newport, & Aslin 1996) and so connection between a phonotactic contribution to parsing and productivity of morphology is also expected. Using their previous findings on the contribution of type frequency for decomposed affixes and the link to productive affix use, they add in a factor for the phonotactic probability of the juncture between the stem and the affix. For example, for the plural *degrees* the juncture is within the /iz/ sequence. Given that this unit is a common rhyme it would not prompt a perceiver to posit a juncture. On the other hand, in *government*, the /nm/ sequence is relatively infrequent word internally and so would likely contribute to parsing. The addition of phonotactic probability at the juncture provides additional predictive power to the model of morphological productivity as a function of parsability. The effect for phonotactic probability is smaller than the effect of relative word and stem frequency but still provides a significant contribution.

The stochastic analyses of morphological productivity presented here led to an additional finding. For an affix that appears in many high token frequency words, the affix is unlikely to be parsed. The type frequency of decomposed forms is not independent from the token frequency of morphologically complex words containing the affix in a way that creates an asymmetry in affix productivity. However, the use of that affix in a variety of high frequency words means that it is useful in communication. If one were to examine a corpus of English there would be two types of frequently appearing affixes in the corpus. Affixes that have a high type frequency that are used with many different stems are likely to be productive. Affixes with high token frequency that appear many times in the same derived forms in corpus are useful. The status of these affixes is likely to change diachronically, however. A truly productive affix that is used in a variety of forms is likely to be stable as a morpheme and used to coin new words. An affix that is useful is less likely to be parsed and so unlikely to be used in new words and is more likely to be subject to semantic drift. Over time, an affix that originated in multiple useful forms may drift enough semantically in different contexts that it becomes opaque.

20.2.3 Morphophonological rules

Lexical or exemplar influences in phonological patterns lie somewhere in between the phonotactic and morphological levels of representation. Phonological rules in generative grammar apply on the basis of symbolically represented criteria. Cases where phonological rules inconsistently apply or where there is variation in the outcome are problematic with a purely symbolic representation. The best studied case of gradience in phonological patterns comes from lexical patterns of place dissimilation in Arabic known as the Obligatory Contour Principle for Place of articulation (OCP-Place as in McCarthy 1994). The Arabic root lexicon is hypothesized to consist of consonant sequences, with vowels and other consonant morphemes interleaved according to morphological templates. At an abstract level, then, the phonotactics of Arabic roots are determined by the permissible and impermissible consonant sequences. Typical roots consist of three consonants. Within a root, more than one consonant at the same place of articulation is generally avoided. However, there are a number of exceptions to this generalization. First, forms where a root appears to contain an identical second and third consonant are common. These forms have been analyzed as representationally consisting of two consonants where the incomplete template is then completed by duplication of the second consonant. Among forms with no identical consonants, consonants at the same place of articulation are more common in first and third position than either first and second or second and third position. Segmental distance in the root affects the restriction on co-occurrence. Finally, for consonants at the same place of articulation, phonological similarity along the other featural dimensions affect co-occurrence. In other words, for consonants at the same place of articulation, their likelihood of co-occurrence is inversely related to their similarity (Frisch, Pierrehumbert, & Broe 2004).

The similarity-based analysis of consonant co-occurrence in Arabic results in a gradient phonological analysis of phonotactics. Similar gradient patterns for place dissimilation have been found in other languages (Frisch, Pierrehumbert, & Broe 2004; Coetzee & Pater 2006) and an analogous similarity pattern has been found for voicing dissimilation (MacEachern 1999). The combination of gradient effects across the lexicon and a similarity-based category structure for the restrictions on segment combinations is completely compatible with the tenets of an exemplar model of phonology. A behavioral experiment with well-formedness judgments for novel verb forms with Jordanian Arabic speakers found that the gradient similarity restrictions of OCP-Place in Arabic are applied to novel forms (Frisch & Zawaydeh 2001).

In addition to examining OCP-Place effects, Frisch & Zawaydeh (2001) provide evidence that multiple influences interact by showing that differences in native speaker ratings of nonce words depend on the violation of the phonotactic constraint, the presence or absence of particular consonant pairs in the lexicon, and the presence of specific similar word forms in the lexicon.

Other phonological phenomena with quantitative patterns are also compatible with an exemplar approach to phonological theory. Zuraw (2008) examines Tagalog nasal substitution where a nasal in the prefix and obstruent initial stem are collapsed into a single combined segment that preserves nasality and the obstruent place of articulation. The rule is inconsistent but more likely to apply to words with labial articulation or with voiceless obstruent onsets. Native speakers productively apply nasal substitution but at a lower rate than it is found across the lexicon. Nasal substitution has also applied to adapted loan words. Zuraw's account of the reality of nasal substitution patterns in the lexicon and in speaker behavior is a dual route model for phonological derivation where derived forms of known words are listed lexically and a phonological process is available to apply to novel words. While the earlier discussion of morphology did not address morphophonological alternations, the analysis of Tagalog nasal substitution is fully compatible with a morphological system where some affixes are parsed and other morphologically complex forms are lexically stored.

Anttila (2002) examines stem final vowel alternations in Finnish morphophonology and finds a striking set of quantitative patterns. In these data, a stem final vowel /a/ with an /i/ suffix can either change to a mid vowel (e.g. /kana-i-ssa/ → [kano-i-ssa]) or delete (e.g. /muna-i-ssa/ → [mun-i-ssa]). Anttila demonstrates that there is phonological conditioning in the process involving the vowel preceding the stem final vowel (/a/ vs /u/ in these examples), the place of articulation of the consonant preceding the vowel (alveolar in both of these examples), and the lexical category of the stem (as noun or adjective). Putting these factors together results in a multidimensional graded category structure where a noun with a preceding /i/ vowel and velar consonant is the prototypical case of vowel height change while an adjective with a preceding /o/ vowel and labial consonant is the prototypical case of vowel deletion. The morphological conditioning is the weakest dimension and is revealed only when the phonological factors do not dominate (e.g. with a preceding /e/ vowel and coronal consonant where variation is common).

20.2.4 Exemplar theories in this chapter

For the purposes of this review, the notion of exemplar theories of phonology is interpreted loosely and so will consider theories based on episodic memory traces (exemplars in the strictest sense) as well as theories that incorporate distributional information such as phonotactic probabilities or neighborhood effects at the more abstract segmental or lexical level. Any notion of an exemplar model of phonology that excludes phonetic data or the processing of phonological information as part of speech production or speech perception would be a model over abstract categories and therefore no longer have access to the details of episodic memory representations (Pierrehumbert 2016). As a result, any model with this level of abstraction in cognitive psychology would really be a considered prototype model rather than an exemplar model (Nosofsky 1992). In phonological discussions of exemplar theories, the distinction between an exemplar and a prototype model is not necessarily maintained. In general, prototype models have the same characteristics of frequency effects and category typicality as exemplar models, so many of the phenomena discussed in this chapter can be accounted for by either model. Some of the data already discussed, for example the phonetic details in frequency-based reduction, can be captured with individual prototype productions for each lexical item. Exemplar representations of individual tokens of production are not needed. By contrast, it would be more difficult to explain the influence of talker identify on word recognition memory without reference to individual exemplars.

20.3 Interfaces with exemplar theory

20.3.1 Explanation of reduction

Bybee's original hypothesis that word usage leads to reduction has spawned an enormous amount of recent research investigating in more detail how words may vary in their word-specific phonetics (Pierrehumbert 2002). Reduction has been investigated for a variety of lexical and functional characteristics and this investigation has been greatly facilitated by the availability of speech corpora that can be practically analyzed on a personal computer. A few examples are covered here.

Bell, Brenier, Gregory, Girande, & Jurafsky (2009) examined a corpus of conversational speech and found that previous mention, high token frequency, and predictability from context each contribute to reducing word duration, despite their intercorrelation. These effects were primarily limited to durations of content words. Function words were found to be shorter than content words, but less influenced by these usage factors. Bell et al. propose a lexical access account that distinguishes access to function words from content words. Rather than concede a special access process is required, it would be worthwhile to examine whether other aspects of the function/content distinction related to position in syntactic or prosodic structure might be at play. In a related study, Gahl (2008) examined homophone word pairs in the Switchboard corpus and found that the same phonological lemma was produced with different durations depending on the frequency of the intended word (e.g. longer low frequency *knead* vs shorter high frequency *need*).

Extending beyond the study of acoustic duration, Ernestus, Lahey, Verhees, & Baayen (2006) found in a corpus of read speech in Dutch that higher frequency words were more likely to show partial or complete voicing assimilation in comparison to low frequency words. This study provides an example of reduction for high frequency words that is distinct from shorter overall duration. In fact, their results are compatible with a shorter duration for the articulatory plan, as cluster duration and voicing were reduced but release noise was increased, suggesting not just a reduction in segment duration but overlap between the segment offset and vowel onset.

Jaeger (2010) examines the use (or absence) of an optional complementizer *that* in a portion of the Switchboard corpus that has been syntactically parsed as part of the Penn Treebank. The use of complementizer that can clarify the intended message for the listener. For example:

My boss confirmed we . . . (ambiguous between NP complement and clause complement)
My boss confirmed that we . . . (clause complement only)
My boss thinks I . . . (clause complement only, redundantly marked)
My boss thinks that I . . . (clause complement only)

Depending on the verb, inclusion of *that* can make the upcoming syntactic structure unambiguous or redundant. Jaeger finds mention of optional *that* is more common when it facilitates understanding without being redundant.

In a case study of lexical frequency effects on variable subject personal pronoun expression in Spanish, Erker & Guy (2012) examine verb frequency and pronoun use in a spoken corpus of Spanish. The use of corpus data supports previous analysis of variable subject pronoun expression based on semantic and syntactic properties of the verb as well as the discourse context. However, the corpus data also reveals that these factors only play a role for high frequency verbs. For low frequency verbs these factors are minimized or absent. In this case the frequency effect is seen only in interaction with other factors that influence expression.

The variety of influences of usage frequency on speech production highlight the importance of a phonological theory in which sub-segmental or phonetic detail is encoded by perceivers. The experimental findings by Goldinger (2000) that a perceiver's subsequent productions are then influenced by these details demonstrates a production–perception feedback loop in adult speakers that would help members of a speech community converge on comparable representations of language sound structure even when the patterns are not course grained.

20.3.2 Social variation

The study of sociophonetics has also undergone rapid expansion recently. An exemplar theory of phonology also readily accommodates sociophonetic cues into processes of perception, production, and the phonological patterns of dialects. Influences of experience on perception of ambiguous phonetic information have shown that the relations between acoustic information and phonological category are influenced by sociolinguistic factors like dialect (Hay, Nolan, & Drager 2006). Hay, Nolan, & Drager (2006) found that perception of vowel categories that vary in different dialects could be primed by non-linguistic information suggesting the likely dialect of the speaker. An exemplar theory of sociophonetic knowledge requires adequate experience with the relevant phonological information. Sanchez, Hay, & Nilson (2015) demonstrated the effect of experience on contextual influences for dialect. When New Zealand English speakers were primed with stereotypically Australian words, their productions shifted toward Australian English norms, but only for those participants with significant experience with Australian English. A shift toward Australian English norms was also found in a corpus study for items in the corpus discussing Australia-related topics.

The learning of statistical distributions in categories extends to gender, age, and social categories. Johnson, Strand, & D'Imperio (1999) demonstrated a similar priming effect on perception based on gender. They found that a visual image of a male or female talker influenced the location of the boundary between /ʌ/ and /ʊ/ in a synthesized continuum. Drager (2011) found that the perception of vowels currently undergoing a historical chain shift was influenced by providing the listener with information about the speaker's age. In a study by Kim (2016), Korean words that are stereotypically associated with older or younger speakers were better recognized by participants when the talker's age matched the word's stereotypical category.

MacFarlane & Hay (2015) discuss an extension of sociophonetics into the broader realm of social psychology. Given that exemplar theories are not specific to phonology, or even to linguistics, there is no reason that the episodic details of experience need to be limited to linguistic or even communicative factors. Under a modular approach to linguistic knowledge there is no expectation that the study of sociophonetics and the study of social psychology would find common ground. The ongoing development of sociophonetic research can continue to test the adequacy of an exemplar theory of language and cognition. However, as broader domains are examined, there can be a challenge to achieve interpretable results. In the cognitive psychology literature on categorization and reasoning there appears to be considerable flexibility in how categories are formed and what dimensions are salient depending, for example, on cultural background and expertise (Medin, Ross, Atran, Burnett, & Blok 2002).

20.3.3 Distribution learning of categories

Studies of infant exposure to patterns in sound categories has provided support for the exemplar, prototype, or statistical approach to the development of phonological categories. When infants are exposed to a variety of related sounds, their learning is influenced by statistical distributions (Saffran, Newport, & Aslin 1996). Further examination of how statistical information is used to form categories has found that distributional information, such as the presence or absence of modes in the statistical distribution of sounds, influences the placement or lack of placement of category boundaries (Maye, Werker, & Gerken 2002). However, a purely distributional learning system will run into difficulties when larger and more realistic data sets are considered. Pierrehumbert (2016) presents a distribution of F3 formant values between /r/ and /t/ allophones in American English males and females. The mode of the distribution for female /r/ lies in roughly the same location as the optimal decision boundary for /r/ vs /t/ for males. These studies are particularly insightful because the experimental stimuli are messy. Individual experimental items are ambiguous because the statistical distributions of the categories overlap. Since overlap in phonological categories is typical along both spectral and temporal dimensions when variation in talker and speech rate is considered, it is important to study how categories can be learned in the face of such variation (Pearl & Goldwater 2016). Fundamental concepts in exemplar models such as similarity in a multidimensional similarity space provide an account of categorization for this kind of messy data. For the language learner, it is necessary to be able to posit multidimensional categorizations and to simultaneously learn the distributions of exemplars and their categorizations (Munson, Edwards, & Beckman 2011).

Pierrehumbert (2016) makes a case for both exemplar and symbolic representations of phonological knowledge as part of the knowledge of language sound structure, resulting in a hybrid exemplar model. The exemplar component of the representation refers to word-specific phonetics, which includes all of the patterns already reviewed involving indexical information and usage frequency effects, predictability effects, and sociophonetic knowledge. These would be related to episodic information in cognitive psychology (Tulving 2002). The other type of representation is abstract categories of the type familiar to generative phonological theory. While not a highlight of the current overview, this type of abstract category is a necessary part of many of the analyses, especially above the phonetic level. While there are statistical distribution properties in the phonotactic patterns of Arabic, for example, those patterns only become apparent when they are examined over abstract roots, and further when root patterns involving identity are handled by a separate mechanism (Berent & Shimron 1997).

Taken together, the abstract categories familiar to generative phonological theory and the exemplar representations that encode word-specific phonetics provide a bridge from phonology to phonetics, sociophonetics, and also to historical language change. Wedel (2006) models some basic properties of evolutionary change in linguistic categories such as natural selection using an exemplar theory. Simulations that utilize the exemplar-based model of the loop between category perception and production crucially require some of the specific assumptions of an exemplar model, such as recency effects. Given the study of the effects of principles such as functional load and contrast in exemplar models of variation, examining natural selection as a principle of historical change in linguistics becomes plausible. Folded together with the utility of exemplar models in integrating phonological patterns with language use, considerable progress has been made toward a comprehensive model of language sound structure incorporating competence, performance, synchrony, and diachrony.

20.4 Conclusions

Overall, exemplar approaches have had a lot of appeal for researchers trying to integrate phonetics and phonology or trying to integrate phonology with other domains that make use of phonology such as psycholinguistics, sociolinguistics, and language acquisition. As these theories have been developed and refined through experimental hypothesis testing and modeling, the field of phonology as a whole has learned a great deal about both the details of phonological/phonetic representation and the necessity for generalization, overarching structure, or abstraction over episodes of memory. On the one hand, phonology has been shown to be closer to traditional phonetics, but on the other hand, robust evidence for the psychological reality of phonological categories themselves has been found.

Exemplar approaches integrate a variety of details beyond a symbolic representation of the phonology. As a result of the inclusion of both phonological and other linguistic information in the representation and processing of language, exemplar theories have had a lot of appeal to researchers trying to integrate phonetics and phonology or trying to integrate phonology with other domains that make use of phonology such as psycholinguistics, sociolinguistics, language acquisition, and language change. While the focus in this handbook is on phonology, one of the strengths of the lexical/exemplar-based approach to phonological patterns is that the exemplar mechanism is not at all specific to phonology. It can be applied at any level of linguistic structure. According to Bybee (2006, p. 730):

> A theory based on usage . . . which takes grammar to be the cognitive organization of language experience, can refer to general cognitive abilities: the importance of repetition in the entrenchment of neuromotor patterns, the use of similarity in categorization, and the construction of generalizations across similar patterns.

References

Albright, A. (2003). A quantitative study of Spanish paradigm gaps. In G. Garding & M. Tsujimura (eds), *WCCFL 22 Proceedings*, 1–14. Somerville, MA: Cascadilla Press.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory*, 20, 1–42.

Bailey, T. & U. Hahn (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44, 568–591.

Bailey, T. & U. Hahn (2005). Phoneme similarity. *Journal of Memory and Language*, 52, 347–370.

Bassok, M. & D. L. Medin (1997). Birds of a feather flock together: Similarity judgments with semantically rich stimuli. *Journal of Memory and Language*, 36, 311–336.

Bell, A., J. M. Brenier, M. Gregory, C. Girande & D. Jurafsky (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60, 92–111.

Berent, I. & J. Shimron (1997). The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition*, 64, 39–72.

Bybee, J. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (eds), *Theoretical morphology*, 119–141. San Diego: Academic Press.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82, 711–733.

Bybee, J. & D. Eddington (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language*, 82, 323–355.

Bybee, J. & E. Pardo (1981). On lexical and morphological conditioning of rules: A nonce-probe experiment with Spanish verbs. *Linguistics*, 19, 937–968.

Carvalho, A. M. (2006). Spanish (s) aspiration as a prestige marker on the Uruguayan-Brazilian border. In C. Mar-Molinero & M. Stewart (eds), *Language variation and change: Historical and contemporary perspectives*, 85–114. Amsterdam: Benjamins.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.

Coetzee, A. W. & J. Pater (2006). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory*, 26, 289–337.

Coleman, J. S. & J. Pierrehumbert (1997). Stochastic phonological grammars and acceptability. *Computational phonology: Third meeting of the ACL special interest group in computational phonology*, 49–56. Somerset, NJ: Association for Computational Linguistics.

Cristia, A., J. Mielke, R. Daland & S. Peperkamp (2013). Constrained generalization of implicitly learned sound patterns. *Journal of Laboratory Phonology*, 4, 259–285.

Dahan, D. & M. R. Brent (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*, 128, 165–185.

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54, 99–121.

Erker, D. & G. R. Guy (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 88, 526–557.

Ernestus, M., M. Lahey, F. Verhees & R. H. Baayen (2006). Lexical frequency and voice assimilation. *The Journal of the Acoustical Society of America*, 120, 1040–1051.

Fanselow, G., C. Féry, M. Schlesewsky & R. Vogel (2006). *Gradience in grammar: Generative perspectives*. Oxford: Oxford University Press.

Frisch, S. A. (2000). Temporally organized lexical representations as phonological units. In M. B. Broe & J. B. Pierrehumbert (eds), *Acquisition and the lexicon: Papers in Laboratory Phonology V*, 283–298. Cambridge: Cambridge University Press.

Frisch, S. A. & M. R. Brea-Spahn (2010). Metalinguistic judgments of phonotactics by monolinguals and bilinguals. *Laboratory Phonology*, 1, 345–360.

Frisch, S. A., N. R. Large & D. B. Pisoni (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, 42, 481–496.

Frisch, S. A., N. R. Large, B. A. Zawaydeh & D. B. Pisoni (2001). Emergent phonotactic generalizations in English and Arabic. In J. Bybee & P. Hopper (eds), *Frequency and the emergence of linguistic structure*, 159–179. Amsterdam: John Benjamins.

Frisch, S. A., J. Pierrehumbert & M. B. Broe (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, 22, 179–228.

Frisch, S. A. & B. A. Zawaydeh (2001). The psychological reality of OCP-Place in Arabic. *Language*, 77, 91–106.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84, 474–496.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, MA: MIT Press.

Goldinger, S. 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, S. D. (2000). The role of perceptual episodes in lexical processing. In A. Cutler, J. M. McQueen & R. Zondervian (eds), *Proceedings of SWAP, spoken word access processes*, 155–159. Nijmegen: Max-Planck-Institute for Psycholinguistics.

Goldrick, M. (2002). Patterns in sound, patterns in mind: Phonological regularities in speech production. PhD dissertation, Johns Hopkins University.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 3–28.

Gries, S. T. (2011). Phonological similarity in multi-word symbolic units. *Cognitive Linguistics*, 22, 491–510.

Harrington, J., S. Palethorpe & C. I. Watson (2000). Does the queen speak the queen's English? *Nature*, 408, 927–928.

Hay, J. (2003). *Causes and consequences of word structure*. New York: Routledge.

Hay, J. & H. Baayen (2002). Parsing and productivity. In G. Booij & J. van Marle (eds), *Yearbook of morphology 2001*, 203–235. Dordrecht: Kluwer Academic Publishers.

Hay, J. & H. Baayen (2003). Phonotactics, parsing and productivity. *Italian Journal of Linguistics*, 15, 99–130.

Hay, J., A. Nolan & K. Drager (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, 23(3), 351–379.

Hillenbrand, J., L. A. Getty, M. J. Clark & K. Wheeler (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61, 23–62.

Johnson, K., E. A. Strand & M. D'Imperio (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, 27, 359–384.

Jurafsky, D., A. Bell, M. Gregory & W. Raymond (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (eds), *Frequency and the emergence of linguistic structure*, 229–254. Amsterdam: Benjamins.

Kim, J. (2016). Perceptual associations between words and speaker age. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1), 18. DOI: http://doi.org/10.5334/labphon.33

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonants clusters. *Journal of Speech and Hearing Research*, 18, 686–706.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 276–312.

MacEachern, M. (1999). *Laryngeal cooccurrence restrictions*. New York: Garland.

MacFarlane, A. E. & J. Hay (2015). Connecting linguistic variation and non-linguistic behaviour. *Linguistics Vanguard*, 1(1), 259–270. DOI: http://dx.doi.org/10.1515/lingvan-2015-1002.

Maye, J., J. F. Werker & L. Gerken (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.

McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed), *Papers in Laboratory Phonology III*, 191–283. Cambridge: Cambridge University Press.

Medin, D. L., N. Ross, S. Atran, R. Burnett & S. Blok (2002). Categorization and reasoning in relation to culture and expertise. In B. Ross (ed), *Psychology of learning and motivation*, 1–41. San Diego, CA: Academic Press.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, 84, 55–71.

Mullennix, J. W. & D. B. Pisoni (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics*, 47, 379–390.

Munson, B., J. Edwards & M. E. Beckman (2011). Phonological representations in language acquisition: Climbing the ladder of abstraction. In A. C. Cohn, C. Fougeron & M. K. Huffman (eds), *Handbook of laboratory phonology*, 288–309. Oxford: Oxford University Press.

Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy & S. M. Kosslyn (eds), *Essays in honor of William K. Estes, vol. 1: From learning theory to connectionist theory, vol. 2: From learning processes to cognitive processes*, 149–167. Hillsdale, NJ: Lawrence Erlbaum.
Nosofsky, R. M. (2014). The generalized context model: An exemplar model of classification. In M. Pothos & A. Wills (eds), *Formal approaches in categorization*, 18–39. New York: Cambridge University Press.
Pearl, L. & S. Goldwater (2016). Statistical learning, inductive bias, and Bayesian inference in language acquisition. In J. Lidz, W. Snyder & J. Pater (eds), *Oxford handbook of developmental linguistics*, 664–695. Oxford: Oxford University Press.
Peterson, G. E. & H. L. Barney (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175–184.
Pierrehumbert, J. (1994). Syllable structure and word structure: A study of triconsonantal clusters in English. In P. Keating (ed), *Papers in Laboratory Phonology III*, 168–188. Cambridge: Cambridge University Press.
Pierrehumbert, J. (2001). Exemplar dynamics, word frequency, lenition, and contrast. In J. Bybee & P. Hopper (eds), *Frequency effects and the emergence of linguistic structure*, 137–157. Amsterdam: John Benjamins.
Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (eds), *Laboratory Phonology VII*, 101–139. Berlin: de Gruyter.
Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics*, 2, 33–52.
Pluymaekers, M., M. Ernestus & R. H. Baayen (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118, 2561–2569.
Rasch, B. & J. Born (2013). About sleep's role in memory. *Physiological Reviews*, 93, 681–766.
Rosch, E. (2003). Prototype classification and logical classification: The two systems. In E. K. Scholnick (ed), *New trends in conceptual representation: Challenges to Piaget's theory?* 73–86. Hillsdale, NJ: Lawrence Erlbaum Associates.
Saffran, J. R., E. L. Newport & R. N. Aslin (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.
Sanchez, K., J. Hay & E. Nilson (2015). Contextual activation of Australia can affect New Zealanders' vowel productions. *Journal of Phonetics*, 48, 76–95.
Skousen, R. (1989). *Analogical modeling of language*. Dordrecht: Kluwer Academic Publishers.
Stemberger, J. P. & B. MacWhinney (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, 14, 17–26.
Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3–46.
Stewart, N. & G. D. A. Brown (2005). Similarity and dissimilarity as evidence in perceptual categorization. *Journal of Mathematical Psychology*, 49, 403–409.
Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53, 1–25.
Wang, H. S. & B. L. Derwing (1994). Some vowel schemas in three English morphological classes: Experimental evidence. *In honor of William S-Y. Wang: Interdisciplinary studies on language and language change*, 561–575. Taipei: Pyramid Press.
Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23, 247–274.
Zuraw, K. (2008). A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language & Linguistic Theory*, 28, 417–472.

21
Algebraic phonology

Iris Berent

21.1 Introduction

Every language – spoken or signed – constructs words from patterns of meaningless elements. English speakers, for instance, *brag* and *speak* but they do not *rbag* or *kspea.* And, as of late, they *blog* and *shvitz* (but not *lbog* or *vshitz*). The existence of such generalizations is, of course, well known. But precisely what mechanisms allow human minds and brains to form phonological patterns and extend them to novel forms is a matter of debate. The controversy concerns not the precise formulation of any particular grammatical rule or constraint (e.g., is nasal place assimilation best explained by Rule1 or Rule2). Rather, the question is whether the grammar – a generative system distinct from the lexicon – exists, and if so, what are the computational properties of this mind/brain faculty.

This chapter considers the hypothesis that the phonological grammar is an *algebraic system*. I contrast the algebraic hypothesis with an alternative view, outlined by associative, exemplar-based accounts of cognition. Both accounts can distinguish between patterns that are attested in the language and unattested patterns. Moreover, both accounts can generalize phonological knowledge to new forms. These accounts differ, however, with respect to the mechanisms they propose to support phonological knowledge, and indeed, mental computations, generally.

The contrast between these views has direct implications for the descriptive and explanatory adequacy of phonological theories. At stake is whether the primitives and combinatorial principles assumed by the generative tradition play a causal role in mental computations. The chapter presents the algebraic hypothesis (as distinct from the associative views) and considers some specific empirical tests that allow one to adjudicate between these alternatives. I conclude by evaluating the algebraic hypothesis against challenges raised by phonetically grounded explanations, and considering the role of exemplar-based knowledge in the phonological system. I suggest that algebraic grammatical rules are at the core of the phonological grammar, and they cannot be subsumed by knowledge of either the statistical structure of exemplars or their phonetic detail. Nonetheless, exemplar-based knowledge (both statistical and phonetic) and rule-based principles are distinct, independent sources of phonological productivity. A full account of phonological competence thus requires a dual-route account that encompasses both algebraic mechanisms and exemplar-based generalizations.

21.2 Does a phonological grammar exist?

To render the discussion as concrete as possible, let us start with a specific example, using a familiar case from English phonology. English requires adjacent obstruent consonants to agree on voicing. Accordingly, tautomorphemic obstruent clusters share the same voicing value ((1)a), whereas those formed productively (e.g., as a result of plural inflection, (1)b–(1)c) are adjusted to avoid a voicing clash.

(1) English voicing agreement: familiar words
- a. Tautomorphemic cluster
  - i. soft, act
  - ii. spot, stock
- b. Plural suffix
  - i. /z/ suffix: Dogs, beds, lads
  - ii. /s/ suffix: Docks, bets, cats
- c. Generalizations
  - i. /z/ suffix: Rogs, neds, rads,
  - ii. /s/ suffix: Dacks, fets, couts

For generative linguistics, such observations immediately suggest the workings of a grammatical principle that enforces voicing agreement (e.g., (2)). The specific formulations of this principle can markedly vary across generative schools. *The Sound Pattern of English* (SPE, Chomsky and Halle, 1968) tradition assumes rules/procedures that convert inputs to outputs, whereas within the framework of Optimality Theory, these generalizations are captured by constraints on outputs (Prince and Smolensky, 1993/2004). Nonetheless, both accounts assume that generalizations are formed by the grammar.

(2) Grammatical accounts of voicing agreement
- a. An SPE-type plural rule (Shibatani, 1973)

  IF: [–sonorant, αvoiced] [–sonorant]##

  ↓ ↓

  THEN: [αvoiced] [αvoiced]

- b. Optimality theoretic (OT) constraint (Lombardi, 1999): obstruent clusters should agree in voicing

What is the evidence that phenomena such as (1) are the product of grammatical rules/constraints? In linguistics, such inferences have been traditionally supported by the observation that the relevant knowledge is systematic (e.g., it applies to any member of the obstruent class) and productive – it applies not only to familiar forms (e.g., in (1)a–(1)b) but also to novel ones (e.g., in (1)c). But this line of reasoning has been called into question by exemplar-based associative models of language (e.g., Rumelhart and McClelland, 1986; Coleman and Pierrehumbert, 1997; Harm and Seidenberg, 1999; McClelland and Patterson, 2002; Joanisse and Seidenberg, 2003; Frisch et al., 2004; Bybee and McClelland, 2005; Hahn and Bailey, 2005; Cole, 2009).

In this view, generalizations are guided not by grammatical principles but rather by the associations between lexical exemplars or their parts (i.e., the statistical structure of the lexicon). The preference for *rogz* (with two voiced consonants) to *rogs* (with a voiced-voiceless sequence) is not the product of voicing agreement per se. Rather, the dislike of *rogs* reflects the fact that voiced-voiceless sequences rarely occur in the English lexicon, whereas voiced-voiced (and voiceless-voiceless) sequences are quite frequent.

(3) Statistical restrictions on feature co-occurrence
*[voiced][voiced]
*[voiceless][voiceless]

Like the plethora of generative grammatical theories, associative exemplar-based models differ widely with respect to the level of abstraction they encode. Inductive grammars (e.g., Coleman and Pierrehumbert, 1997; Frisch et al., 2004; Hayes and Wilson, 2008; Albright, 2009; Adriaans and Kager, 2010) extract from the lexicon constraints on feature co-occurrence, such as a ban on adjacent voiced (and voiceless) features as in (3). Although these constraints are based on lexical associations, they are proposed as an account of a grammar, and their form appears quite similar to the SPE and classical OT grammatical statements (e.g., (2)). A far more radical alternative to the generative tradition is presented by connectionist models (e.g., Rumelhart and McClelland, 1986; Harm and Seidenberg, 1999; McClelland and Patterson, 2002; Joanisse and Seidenberg, 2003; Monaghan and Shillcock, 2003; Bybee and McClelland, 2005; Hahn and Bailey, 2005). These models eschew the distinction between the lexicon and the grammar altogether. In this approach, the familiar *dogz* and novel *rogz* acquire precisely the same kind of representation – a vector of activation among nodes (i.e., representational units) that encodes phonological features. The preference for the voiced *rogz* over the voiceless *rogs* reflects the higher statistical frequency of this combination in a speaker's linguistic experience – knowledge that is essentially the same as the one predicting the familiarity with *dog* (but not *rog*) as an English word. In both cases, frequent statistical patterns are preferred. Crucially, the generalization of the voicing pattern to novel forms does not imply any grammatical principle at all. Likewise, these models eliminate the encoding of most grammatical categories – it includes no representations for syllables, onsets, feet or morphemes.

These observations raise profound questions concerning the theoretical enterprise of phonological theory, generally, and the role of the grammar, specifically. Within the generative tradition, the grammar is a synonym to productivity, as all aspects of systematic linguistic productivity are thought to rely on grammatical principles. But if productivity can be attained by a model that eliminates the grammar altogether, then what is the role of rules and constraints in the language system? Does the language system effectively encode a constraint on voicing agreement (as in (2)), or does it only track the statistical structure of the lexicon? More generally, does the grammar form a distinct system of the mind/brain, or is it merely a descriptive label for lexical mechanisms that rely on associative, statistical learning?

21.3 The algebraic account of the grammar

To begin exploring the role of the grammar, we must first ask whether grammatical rules and constraints, as in (2), share any common properties, and whether those properties are distinct from statistical principles, as in (3). To the extent such properties are found, we can next seek empirical evidence that could adjudicate between these proposals. These questions require that we take a closer look at the anatomy of grammatical principles.

Viewed in passing, the principles in (2)a–b indeed look very different from each other, and quite different from the statement in (3). But a closer inspection reveals an important distinction. The two statements in (2) explicitly appeal to the *identity* of the two voicing features. Statement (2)a generates outputs whose voicing feature has identical values by copying (i.e., reduplicating) the voicing value of the first segment. The constraint in (2)b likewise enforces voicing identity by requiring the two obstruents to agree on their voicing feature. Obviously, the two statements in (2)a–b follow different means to enforce voicing identity, but the goal – voicing identity – is shared. In contrast, the statistical statement in (3) captures feature co-occurrence: it indicates that these two features do not follow a given sequence. However, there is nothing in the statement of this constraint that appeals to their identity. And indeed, a constraint such as *[voiced][voiced] could apply just as well to combinations of nonidentical features (e.g., *[sonorant][obstruent]).

But identity and feature co-occurrence impose different computational demands on learners. The identity function links a member *i* of a given category (e.g., voicing) to another such *i member* ($\alpha_i\ \alpha_i$). To ensure the identity of these two instantiations, it is thus necessary to bind them by an algebraic *variable* (α). For example, if the first occurrence of α_i (in $\alpha_i\ \alpha_i$) is voiced, then a voiced feature will also occupy the second position. The binding of these two occurrences by the variable α enforces their identity in a manner akin to the role of variables in arithmetic expressions (e.g., the same number must instantiate all occurrences of X in $2X^2 + X + X = y$).

The statements in (2)a–b both include operations over variables: (2)a copies a variable (α); statement (2)b enforces agreement among any two instances of that variable. Moreover, the class that is restricted by the identity rule/constraint is defined solely by a structural property that is independent of specific exemplar members. Here, the relevant class concerns voicing (other classes include "onset", "coda", "syllable", etc.). Because the conditions for membership in the class are structural (i.e., whether a feature marks voicing), all exemplars are equally good members of the category (i.e., they are all representative of the category as a whole). Thus, *p* is as good as a voiceless consonant as *θ*, despite the fact that *p* is more frequent than *θ* in English. As such, the voicing category forms an *equivalence class*.

In fact, this class (e.g., the class of all voiceless consonants) is *open-ended* – it includes not only the (small) set of voiceless features that actually occur in the lexicon, but also potential ones (e.g., ones that can be introduced due to borrowings). And since the identity rule/constraint is defined only as an operation over variables ((4)a), and since these variables stand for equivalence classes ((4)b), these principles are guaranteed to apply across the board, to any class member – actual or potential.

Operations over variables, then, have the potential to apply to all members of a class, and consequently they generalize across the board, to all novel members. As such, systems that encode statements such as (2)a–b support unbounded infinity ((4)c) – the celebrated feature of the grammar that is typically only considered in the context of syntax (Chomsky, 1972). In practice, however, many phonological categories are quite limited and small, so it is difficult to compare the predictions of algebraic principles with mechanisms of statistical learning that track feature co-occurrence from lexical instances. In the present case of voicing identity, the class of voicing only includes two members (voiced and voiceless), so the contrast between grammatical principles in (2) and the exemplar-based generalization in (3) is difficult to establish. Nonetheless, the distinction between identity restrictions in (2) and the statistical statement in (3) is not moot. Indeed, identity restrictions have been proposed for numerous aspects of phonology in categories (e.g., Leben, 1973; McCarthy, 1979; Yip, 1988; McCarthy, 1994; Suzuki, 1998; Rose and Walker, 2004). Similarly, reduplication (the mirror image of identity restriction) is widely documented in the grammar, and it can acquire purely phonological functions (Alderete et al., 1999; Bat-El, 2006; Inkelas, 2008). Such cases, as we shall shortly see, can potentially extend even to feature values that are unattested in the language. It is those cases that reveal the contrast between the operations over variables and exemplar-based statements in the clearest way.

Summarizing, then, grammatical restrictions such as (2)a–b exhibit three key properties that are characteristic of a class of computational systems known as *algebraic* (also known as symbolic) accounts of cognition (see (4)).

(4) Some properties of algebraic rules
  a. Algebraic rules operate on variables.
  b. Algebraic rules appeal to equivalence classes that are potentially open-ended.
  c. Algebraic rules support unbounded generalizations.

Note that algebraic rules (as defined above) encompass both SPE-style rules and OT constraints, and as such, the notion of an algebraic rule should be kept distinct from the narrow technical notion of rules, as employed in linguistics (i.e., as the mapping from inputs to outputs). The algebraic hypothesis is rooted in the computational theory of mind, proposed by Jerry Fodor and Zenon Pylyshyn (Fodor, 1975; Pylyshyn, 1984; Fodor and Pylyshyn, 1988), following the work of Alan Turing (Turing, 1936, 1950). The hypothesis that linguistic rules consist of algebraic operations over variables has been defended by Steven Pinker and Alan Prince in their seminal discussion of the past tense rule (Pinker and Prince, 1988) and a large experimental research program on inflectional morphology (e.g., Pinker, 1991; Prasada and Pinker, 1993; Kim et al., 1994; Marcus et al., 1995; Pinker, 1999). Subsequent research by Paul Smolensky and colleagues (Prince and Smolensky, 1997; Smolensky and Legendre, 2006) has shown that the representation of syllable structure is the product of algebraic operations that can be implemented in certain connectionist networks. The role of algebraic mechanisms in phonology has been further investigated in a large experimental research program by Iris Berent and colleagues (e.g., Berent and Shimron, 1997; Berent et al., 2001a; Berent et al., 2002; Berent et al., 2007) and summarized in Berent (2013a, 2013b). We now move to review some of the key findings supporting this hypothesis.

21.4 Experimental tests of the algebraic hypothesis

Algebraic and exemplar-based approaches make distinct predictions regarding the nature of phonological competence. The algebraic hypothesis predicts that phonological knowledge encodes abstract relations by representing the constituent structure of categories using mental variables. Instance-based accounts, by contrast, assume a much leaner computational machinery – its precise nature and overlap with the algebraic proposal varies across models. Some statistical learning models (e.g., the Maxent model; Hayes and Wilson, 2008) include equivalence classes, but not operations over variables (see section 21.5.2 for details); radical connectionist approaches (feedforward networks, simple recurrent networks; Rumelhart and McClelland, 1986; Harm and Seidenberg, 1999; McClelland and Patterson, 2002; Joanisse and Seidenberg, 2003; Monaghan and Shillcock, 2003; Bybee and McClelland, 2005; Hahn and Bailey, 2005) eliminate the contrast between instances and categories altogether – they include no representation of a syllable, for instance, distinct from specific exemplars (e.g., the chunks *pen* and *cil* in *pencil*). Likewise, these models eliminate variables and operation over variables. Nonetheless, connectionist models have been shown to exhibit phonological knowledge and generalizations that mimic certain aspects of human behavior. These similarities open the door to several distinct interpretations (see (5)).

(5) How can nonalgebraic mechanisms capture phonological generalizations?
  a. Algebraic rules do not exist; phonological knowledge is strictly exemplar-based.
  b. Algebraic rules exist, and they cannot be subsumed by nonalgebraic connectionist mechanisms.
  c. Algebraic rules exist, but they are an "emerging property" of nonalgebraic connectionist systems.

One possibility is that phonological knowledge is strictly exemplar-based, rather than algebraic. Alternatively, phonological knowledge might effectively include an algebraic grammatical component that is distinct from exemplar-based associations, but because the ranking of grammatical constraints optimize harmony (Prince and Smolensky, 1993/2004), grammatical and lexically based preferences often converge on their prediction. Finally, there exists the possibility that phonological principles are algebraic (e.g., they encode equivalence classes and operate on variables), but these properties emerge spontaneously in connectionist systems that are initially designed to only encode feature-co-occurrence.

To adjudicate between these approaches, we first ask whether an account of human phonological competence requires algebraic mechanisms. Specifically we ask whether phonological knowledge includes abstract equivalence classes, and whether such classes form the basis of algebraic rules. Whether such properties can emerge in connectionist systems is a question we revisit in the next section.

21.4.1 The role of equivalence classes

Algebraic accounts of phonology include abstract categories such as "syllable", "onset" and "foot". Each such category (e.g., syllable) appeals to a class of exemplars (e.g., the English syllables *ped*, *a*, *act*), yet category and its members are distinct: the notion of a "syllable" is represented separately from its instances, and it cannot be subsumed by them. Moreover, categories such as "syllable" form equivalence classes, as their members are all treated alike with respect to relevant generalizations, irrespective of their frequency in the language.

Such categories, however, are eliminated from exemplar-based connectionist models. An influential paper by Mark Seidenberg (1987) asserted that syllables are not encoded by readers. To explain why people treat words like *anvil* as disyllabic, Seidenberg appeals to the statistical structure of sub-syllabic chunks (e.g., *an* and *vil* in *anvil*). And indeed, syllable boundary is often marked by a frequency trough. For example, consider the statistical co-occurrence of bigrams (two-letter combinations) in the printed word *anvil. Anvil* includes three bigrams: *an*, *nv*, *vi*, *il*. These bigrams differ markedly on their frequency in the English language (for statistical information, see (6)). While the bigrams that open each syllable (*an* and *vi*) are quite frequent, the bigram that stands the syllable boundary (*nv*) is rare. The bigram trough (relative to the surrounding units) provides critical information that could help parse the words into two chunks (*an-vil*). Crucially, this chunking is based not on the encoding of syllables but rather on the statistical properties of instances or chunks.

(6) Bigram troughs in *anvil* vs. *igloo*
Note: All bigram counts are based on five-letter English words. Count provides the number of words that include a given bigram; Frequency provides the summed frequency of those words; all calculations are position sensitive and counts are per million; they are extracted from Solso and Juel (1980) based on Kucera and Francis (1967).

a. anvil

String	COUNT	FREQENCY
an	23	289
nv	***2***	***5***
vi	24	324
il	22	738

b. igloo

String	COUNT	FREQENCY
ig	0	0
gl	***1***	***2***
lo	18	423
oo	2	4

In view of such observations, one wonders what strategies help readers chunk words into syllable-like units – do they effectively encode syllables, or do they only track frequency troughs?

To distinguish between these possibilities, Brenda Rapp (1992) examined people's sensitivity to the syllable structure of two types of words. In one type, syllable boundary was marked by a bigram trough (e.g., as in *anvil*). A second type of words lacked such bigram trough information. For example, in the word *igloo*, the critical *gl* bigram (i.e., the syllable trough) is slightly more frequent than the initial bigram *ig*. If syllable structure is an artifact of bigram trough, then its effects should be limited to the first (*anvil* type) but not second (*igloo* type). But the results suggested that people encoded the syllable boundary in both types. Moreover, the effect of the syllable emerged automatically, despite the fact that the experimental task does require attention to syllable structure (participants were asked to name the color of letters that were presented subliminally). We should note that statistical information also plays an important role that can sometimes greatly attenuate the effect of syllable structure (Doignon and Zagar, 2005). Nonetheless, these results suggest that syllables are abstract categories, distinct from their members and irreducible to their statistical properties.

21.4.2 The role of algebraic rules

At the center of the algebraic proposal is the hypothesis that phonological knowledge is encoded by a set of algebraic rules, defined as operations over variables. Two critical features distinguish algebraic rules from exemplar-based alternatives. First, algebraic rules are guided by the syntactic structure of mental representations. For example, because identity is enforced by binding variables (e.g., $\alpha_i \alpha_i$), identical elements have complex constituent structure, distinct from nonidentical elements (e.g., $\alpha\ \beta$). In the algebraic proposal, this constituent structure is explicitly represented, and indeed, it has a causal role in guiding mental processes (i.e., *baba* differs from *bada* because the former is assigned the structure of $\alpha_i \alpha_i$). And because this structure is defined over abstract categories that are potentially open-ended, algebraic rules could, in principle, extend across the board. The encoding of constituent structure and across-the-board generalizations are the two hallmarks of algebraic rules. The next two sections evaluate these two predictions.

21.4.2.1 The encoding of constituent structure

The algebraic hypothesis forms part of the computational theory of mind (e.g., Fodor and Pylyshyn, 1988). In this view, mental representations are structured symbols: they have both form (the signifier) and semantic contents (the information that is signified). Moreover, the form of mental representations plays a causal role in mental processes. For example, forms like *bagag* encode the identity of the final consonant (e.g., ABiBi), whereas no such information is present in ABC forms like *bagad* (with three nonidentical consonants). If this explanation is correct, then people should distinguish novel ABB and ABC forms even when these items are matched for their consonant co-occurrence (and are both well-formed in Hebrew). This prediction is borne out by the results of numerous experiments from Hebrew (Berent and Shimron, 1997; Berent et al., 2000; Berent et al., 2001a; Berent, 2002; Berent et al., 2002; Berent and Shimron, 2003; Berent et al., 2006; Berent et al., 2007; Berent et al., 2012a).

Consider, for example, the results from lexical decision experiments (Berent et al., 2001b; Berent et al., 2004; Berent et al., 2007). In these experiments, people are presented with a single stimulus – either an existing Hebrew word, or a novel word-like stimulus – and they are asked to quickly indicate whether this stimulus is a real Hebrew word. Results showed that novel ABB stems (e.g., *bagag*) produce significantly slower "nonword" responses compared to ABC controls (e.g., *bagal*), and these results obtain even when both forms are matched for the co-occurrence of their consonants (using type-frequency measures) and their vocalic patterns.

The difficulty to identify ABB stems as nonwords implies that ABB stems are better-formed than ABC controls – possibly, because their reduplicative structure is relatively unmarked. However, identical elements are also known to be banned in phonology, and indeed, several rating tasks have shown a dispreference for ABB items (Berent and Shimron, 1997; Berent et al., 2001a). These variations suggest that identity might have distinct consequences (i.e., preference vs. dispreference), which are partly task-dependent. Our question here, however, concerns not the consequence of identity but rather its very encoding. The fact that people attend to the structural contrast between ABB and ABC forms suggests that they represent the constituent structure of identical elements. In fact, structure sensitivity is evident quite early in life. An influential set of experiments by Gary Marcus and colleagues has documented the sensitivity to reduplicative structure in 7-month-old infants (Marcus et al., 1999), and subsequent findings by Judit Gervain and colleagues has extended this work to neonates (Gervain et al., 2012). The sensitivity of adults and infants to the constituent structure of phonological forms is in line with the algebraic account.

21.4.2.2 The scope of phonological knowledge

Another test for the representation of rules is presented by the scope of phonological generalizations. The hallmark of an algebraic rule is its capacity to extend generalizations freely – to any novel member of a class. So if phonology is an algebraic system, then phonological generalizations should be in principle unbounded. This prediction, however, is often difficult to evaluate, as the phonological categories attested in a given language (e.g., voicing) are finite and quite small (e.g., voiced vs. voiceless phonemes). But the small size of phonological categories does not necessarily imply that phonological rules are inherently narrow in scope. To evaluate the scope of phonological generalizations, one can ask whether phonological generalizations extend to elements that are unattested in participants' language. The algebraic hypothesis predicts that such generalizations should be possible, but, as we next demonstrate, these predictions do not follow from exemplar-based models.

To contrast between these predictions, consider again the structure of Hebrew stems. As noted above, Hebrew allows forms such as ABB (e.g., *bigeg*) and ABC (e.g., *biges*). Like other Semitic languages, however, Hebrew bans forms such as AAB (e.g., *bibg*), and people generalize this ban to novel forms (Berent and Shimron, 1997; Berent et al., 2000; Berent et al., 2001a; Berent et al., 2002; Berent et al., 2006; Berent et al., 2007; Berent et al., 2012a). Of interest is whether this generalization is explicable by the statistical co-occurrence of AAB elements or their structure. An algebraic rule (*AAB, where A stands for any consonant) should generalize across the board, even when the relevant consonant is not native to the language.

Consider, for example, generalizations to the consonant /θ/. This consonant is not native to Hebrew, and no Hebrew word includes this consonant (e.g., borrowings including /θ/ are routinely transformed to /t/). Moreover, not only is the /θ/ phoneme unattested, but so is its place of articulation feature (the wide value of the tongue tip constriction area feature; Gafos, 1999). Since stems like *kaθaθ* are arguably as infrequent as *θaθak* with respect to both segment and feature co-occurrence, statistical knowledge should offer no help in distinguishing between them. In contrast, if people represent the reduplicative structure by an algebraic rule, then they should readily encode *θaθak* as AAB, hence, as worse-formed than its ABB counterparts, *kaθaθ*. A series of experiments revealed precisely this pattern (Berent et al., 2002). Specifically, *θaθak*-type stems elicited lower acceptability relative to *kaθaθ*-type stems, and this finding obtained across various morphological contexts, even when these stems were heavily prefixed and suffixed. Likewise, *θaθak*-type elicited faster lexical decision time, suggesting that such stems are less word-like than *kaθaθ*-type counterparts. Together, these results demonstrate that people extend the AAB rule across the board – even to novel consonants with novel features.

21.5 Computational tests of the algebraic hypothesis

The results reviewed so far suggest that people generalize their phonological knowledge in line with the algebraic hypothesis – they seem to encode abstract categories, and they are sensitive to the constituent structure of abstract variables even when statistical information is controlled, and indeed, even when no relevant statistical information is available because the relevant elements are entirely unattested in participants' language. But showing that human behavior is *consistent* with algebraic rules does not necessarily demonstrate that such rules are encoded in the mind and brain, and that these mechanisms are the ones that effectively cause the observed behavioral pattern. And indeed, one might wonder whether the generalizations we attribute to algebraic rules could also emerge spontaneously in mechanisms that only encode the statistical co-occurrence of instances (as per (5)c above). Such an outcome would still be consistent with the hypothesis that phonological generalizations are unbounded, but if it were true, such generalizations would not require a grammar. While this question remains the topic of active research, the existing results concerning identity restrictions suggest that this possibility is unlikely. The critical evidence comes from studies that specifically examine the capacity of various computational systems to extend the identity function.

21.5.1 Can connectionist networks give rise to algebraic rules?

One line of evidence is presented by the computational simulations of Gary Marcus (1998, 2001). In this research, Marcus examined the capacity of various popular connectionist networks (e.g., feedforward and simple recurrent networks) to extend the identity function across-the-board – to any member of a class. To address this question, one must first define more precisely what is meant by "across the board" generalizations. Once generalizations are defined, one can proceed to ask whether they are attainable by associative connectionist networks. We will discuss these two questions in turn.

21.5.1.1 Defining the scope of phonological generalizations

To define the scope of (phonological) generalizations, Marcus proposes to compare the representations of familiar and novel items. For the sake of concreteness, let us consider a specific case. In our example, a network is trained on a reduplication rule X→XX (see (7)). To this end, the network is first trained on the reduplication of two training items, *pa* and *ta* (e.g., *pa*→*papa*). After mastering the association between inputs (e.g., *pa*) and outputs (e.g., *papa*), the network is next tested for its ability to generalize this function to novel inputs – either *ba* or *xa*. Although these testing items are not encountered in training, generalization is not necessarily unattainable. Rather, generalization strictly depends on the overlap between the representations of training and generalization items.

(7) Generalizing the reduplication function

Training:	pa→papa
	ta→tata
Generalization:	
	ba→?
	xa→?

Overlap is critical because, in these networks, knowledge is captured as an association between input and output units. For example, the *ba*→*baba* correspondence is represented by association between the input nodes for *ba* and the output (*baba*). But if each such item is represented by a single input, and this node does not form part of the representation of novel items (e.g., *pa*), then training is essentially useless: knowledge that the *pa* input node activates the *papa* output is entirely informative for predicting the desired output for *ba*, so generalizations should fail for both *ba* and *xa*. In contrast, if the same two syllables are represented by features, then the representations of the familiar *pa* and *ta* will now overlap with *ba*; the example in (8) illustrates the overlap on the place of articulation feature. This overlap will allow a network to generalize the reduplication function to *ba*. In the case of *xa*, however, the overlap with training items will be incomplete: training on *pa* presents no guidance on the desired activation for the place feature of *xa*, and consequently, generalization to this item will be far more challenging.

(8) The overlap between the representations of training and generalization items

		Syllables	Features (labial)
Training	pa	+	+
	ta	+	−
Generalizations	ba	−	+
	xa	−	−

Summarizing, then, Marcus proposes to define the scope of generalizations in reference to a network's representational space. Items that can be exhaustively represented within the representational space of training items fall *within* the network's training space; those that cannot be exhaustively described in this fashion fall *outside* the training space. In the above example, *ba* (but not *xa*) falls within the training space of a network that encodes features; for a network that encodes syllables, both items fall outside the training space.

21.5.1.2 Can exemplar models generalize across the board?

Armed with the notion of the training space, we can now define across-the-board generalizations as ones that support generalizations even outside the training space. Algebraic mechanisms, by definition, should generalize either within or beyond the training space. These conclusions now set the stage for returning to our original question of whether across-the-board generalizations are attainable by connectionist networks that lack "innate" operations over variables.

Marcus's results suggest that this is not the case. This is not because the networks categorically failed to generalize. Indeed, novel items that fall within the training space (e.g., *pa*, given *ba* for the feature network) did elicit the desired reduplication output. Crucially, reduplication failed for novel items that fall beyond the training space (e.g., for *xa*). This result suggests that these networks can acquire productive knowledge that supports generalizations. Unlike an algebraic rule, however, this knowledge is limited in scope, as it does not support generalizations across the board.

The experimental results presented in the previous section suggest that such generalizations are readily attained by humans. Unlike these connectionist networks, people generalize outside the training space of their language, as they extend the *AAB rule to novel items with novel feature values. It is important to note that these conclusions are specific to a certain class of connectionist networks, rather than to connectionism generally. Other results suggest that some connectionist networks can represent algebraic rules, provided that the network is equipped with algebraic mechanisms that distinguish categories and instances, and operate on categories as a whole (i.e., operations over variables; for a potential example, see Smolensky and Legendre, 2006). Crucially, absent such "innate" algebraic mechanisms, across-the-board generalizations fail. Thus, the encoding of algebraic rules is necessary to exhibit across-the-board generalizations. This very same conclusion also emerges from a second line of inquiry with the Maxent model.

21.5.2 Algebraic rules in the Maxent model

Maximum Entropy (Maxent) models (Hayes and Wilson, 2008) have been the subject of much recent interest in phonology. These models differ from the connectionist networks studied by Marcus in several important respects. Unlike connectionist networks, Maxent models capture generalizations by inducing grammatical constraints that are encoded separately from lexical instances. However, the original Hayes and Wilson (2008) model nonetheless resembles the above-mentioned connectionist networks inasmuch as it lacked variables. For example, this model can extract the fact that two labials rarely co-occur at a given word position (e.g., by the constraint *[labial]–[labial]), but it has no means to represent the fact that those labials are in fact *identical* (e.g., per the OCP). If operations over variables are, in fact, necessary to capture phonological knowledge, then this model should fail to generalize the *AAB rule across the board.

A series of simulations suggest this is indeed the case (Berent et al., 2012b). These simulations first trained the model on the set of Hebrew consonantal roots (for simplicity, these simulations ignored intermediate vowels). Next, the model was tested on the set of materials used in behavioral experiments. Of interest is whether the model could capture the human preference for novel ABB forms over AAB ones.

Results showed that the model was, in fact, able to generalize. But, once again, generalizations were limited in scope. When test items comprised of native features (i.e., ones represented in the training items), the model exhibited the desired generalizations. But when presented with the critical *θ*-items, whose place of articulation is nonnative to Hebrew (e.g., *θaθak* vs. *kaθaθ*), no preference for ABB items was found (for similar conclusions obtained from another case of identity restriction, see Gallagher, 2013).

To determine whether the model's failure was specifically due to the lack of operations over variables, the model was next fitted with such mechanisms. This modification had a dramatic effect on its performance. Once variables were introduced, the model exhibited the human preference for *kaθaθ* over *θaθak* items. These results converge with the findings from connectionist networks to suggest that generalizations beyond the training space require algebraic operations over variables. Crucially, people freely generalize their phonological knowledge in this fashion. As such, these results suggest that phonological competence includes algebraic rules.

21.6 Conclusions, challenges and extensions

This chapter examined the hypothesis that grammatical phonological generalizations are the product of algebraic mechanisms. To test this hypothesis, we first examined the role of equivalence classes and algebraic rules in two representative case studies. In support of equivalence classes, we showed that readers' sensitivity to the syllable cannot be subsumed by the statistical co-occurrence of letters. We next moved to examine whether speakers encode restrictions on identical elements by means of algebraic rules. To this end, we examined two properties of algebraic rules: structure sensitivity and across-the-board generalizations. We reasoned that if people represent algebraic phonological rules, then they should be sensitive to the structural distinction between ABB and ABC forms, even when their statistical properties are matched. The results of numerous studies are in line with this prediction. A second hallmark of algebraic rules is the capacity to extend generalizations to any member of the relevant class, even when this item has never been encountered in the context of that generalization. To define across-the-board generalizations, we introduced the notion of the training space, and we pointed out that algebraic rules allow learners to generalize beyond their training space. Results from experimental studies show that people do, in fact, freely extend phonological generalizations, and computational simulations confirm that such generalizations are only attainable by mechanisms that encode algebraic rules. Accordingly, the available evidence suggests that the phonological grammar is endowed with powerful algebraic mechanisms, and it is irreducible to the statistical structure of the lexicon.

These conclusions are indeed consistent with a large body of research in formal linguistics, in which equivalence classes (e.g., syllable) and algebraic rules (e.g., OCP, identity by correspondence constraints, reduplication) are tacitly assumed on a routine basis. The algebraic hypothesis explicitly outlines the computational mechanisms necessary to implement these linguistic proposals. But the appeal to algebraic rules would seem to conflict with two other literatures that have gained much traction in phonological research. One such literature documents speakers' productive knowledge of phonological exemplars (e.g., Pierrehumbert, 2001; Bybee and McClelland, 2005). Another challenge to the algebraic hypothesis is presented by the strong links between the phonological and phonetic systems (e.g., Archangeli and Pulleyblank, 1994; Hayes et al., 2004). We now move to consider those challenges. As we next demonstrate, the conflict with the algebraic hypothesis is only apparent. Nonetheless, the undeniable support for exemplar-based knowledge and the strong phonetic–phonology link call for a richer account of phonological competence, in which productive generalizations are the product of multiple mechanisms – both grammatical and nongrammatical. One such proposal is outlined in the final section.

21.6.1 Gradience in phonological preferences

If phonological knowledge consists of a set of algebraic rules that determine well-formedness, then the effect of well-formedness should be evident in phonological behavior. And if phonological well-formedness consists of a binary contrast between well-formed and ill-formed structures, then this distinction should give rise to a categorical distinction in behavior. But decades of research in psycholinguistics have made it clear that phonological preferences are typically gradient. Acceptability judgments, for instance, normally reflect a relative preference for better-formed structures over ill-formed ones; ill-formed structures are rarely banned categorically as "impossible". Moreover, within the class of the "better stimuli", not all members are treated alike. Rather, the recognition of the stimulus and its acceptability is often modulated by exemplar-specific properties, such as frequency and similarity. Familiar words such as *cat* and *dog* are identified more readily than the less familiar *fog* (e.g., Taft, 1979; Balota and Chumbly, 1984). Likewise, similar items (e.g., *gat–cat*) prime each other in perception (e.g., *cat* is recognized more readily when preceded by the similar *gat* relative to the less similar *bat* (e.g., Perfetti et al., 1988; Perfetti and Bell, 1991)), whereas in speech production, they can produce a tongue twister (McCutchen et al., 1991). These strong exemplar-based effects appear to fly in the face of the hypothesis that the grammar encodes equivalence classes that are blind to the distinction between class members. Similarly, the gradient, relative pattern of preferences would seem to counter the view of well-formedness as the product of an algebraic grammatical rule.

But the challenge from exemplar-based models is even stronger. Not only are people exquisitely sensitive to the properties of specific lexical instances, but those properties can even support generalizations. Numerous studies have shown that the acceptability of novel words is predicted, at least in part, by their various measures of their similarity to existing lexical forms (e.g., bigram/biphone measures, neighborhood density, e.g., Luce and Pisoni, 1998; Perea and Carreiras, 1998; Vitevitch et al., 1999; Ziegler et al., 2003; Hahn and Bailey, 2005). In fact, these lexical effects can even counter the effects of grammatical rules. Recall, for example, that Hebrew (and many other Semitic languages) ban AAB type stems (e.g., *sisem*). While AAB forms are clearly underrepresented in the lexicon, these items are not categorically banned. Certain AAB forms are attested in Hebrew (e.g., *mimen* 'financed', *mimesh* 'realized'). And when presented with novel AAB forms that are analogous to those counterexamples (e.g., *mimek*), Hebrew speakers show no dispreference (relative to ABB forms; Berent et al., 2001a). Such observations show that counterexamples are stored in memory, and support generalization (by analogy) to novel forms.

But while the associative effects of exemplars are undeniable, the challenge they present to the algebraic hypothesis is more apparent than real. Indeed, the hypothesis that phonological knowledge includes an algebraic grammar does not imply that phonological knowledge is restricted to the grammar alone. The strong effects of familiarity and similarity suggest that some aspects of phonology are exemplar-based, but these conclusions do not negate the possibility that algebraic rules might have an additional role.

In a similar vein, grammatical well-formedness need not map transparently to acceptability or processing ease. Like any other aspects of performance, the recognition of linguistic stimuli and their acceptability are the product of multiple systems. Grammatical knowledge, if it exists, is one of those systems, but so are lexical association, attention allocation and strategic control in response to specific experimental settings. Accordingly, a binary algebraic distinction between well-formed and ill-formed items should disfavor ill-formed items, but it should not necessarily render these judgments categorical (e.g., as "impossible"). In fact, the algebraic hypothesis does not necessarily require that well-formedness is confined to binary contrasts. While some markedness constraints are binary (e.g., Onset), others take the form of markedness hierarchies (e.g., a >*b >*c) that compute gradient (rather than categorical) harmony functions as their outputs (Prince and Smolensky, 1993/2004; de Lacy, 2006a). In addition, the algebraic hypothesis is perfectly in line with the possibility that (algebraic) grammatical constraints are weighted. For all these reasons, the representation of algebraic rules could well result in gradient performance.

To evaluate the algebraic hypothesis, then, the principal question is not whether people are insensitive to specific exemplars; clearly, they are. Rather, at stake is whether exemplar-specific properties are *sufficient* to capture the full range of phonological knowledge. The previous sections suggest that the answer to this question is firmly negative, as (a) people are sensitive to constituent structure, and use it to generalize their phonological knowledge; (b) phonological generalizations extend beyond the training space, in the absence of relevant statistical exemplar-based knowledge; and (c) such generalizations are demonstrably unattainable in the absence of algebraic rules. Accordingly, a complete theory of phonological knowledge ought to include both associative mechanisms that track the statistical co-occurrence of exemplars along with an algebraic grammatical system.

21.6.2 The phonetic grounding of phonology

Another challenge to the algebraic hypothesis is presented by the grounding of phonological rules in phonetics. These links are documented in dozens of studies, and their discovery is arguably among the most important contributions of modern research in phonology (e.g., Ohala, 1975; Browman and Goldstein, 1989; Archangeli and Pulleyblank, 1994; Steriade, 1997; Hayes et al., 2004). Generally speaking, preferred phonological patterns are ones that optimize perception and facilitate articulatory production. For example, the preference for unmarked syllable shapes (e.g., CV>VC; *blog>lbog*) maximizes coarticulation and benefits speech perception and production (Mattingly, 1981). Similarly, segmental and tonal contrasts are grounded in phonetic constraints (e.g., Flemming, 2004; Zhang, 2004). For example, the inclusion of coronal but not labial voiceless stops in Egyptian Arabic has been attributed to the greater articulatory demands associated with the manipulation of voicelessness in labials (Hayes, 1999; Wright, 2004).

Moving to consider phonological processes, one is immediately struck by the many parallels between phonological and phonetic interactions. Assimilation, for instance, closely mirrors the phonetic processes occurring naturally during coarticulation (e.g., Jun, 2004). Similarly, phonological repair (i.e., in response to markedness violations) typically conspires to obscure the inserted material (Steriade, 2001). For example, the ban on final voiced segments is often met by the devoicing of the final segment (e.g., *tab*→*tap*), but not its nasalization (e.g., *tab*→*tam*), as devoicing contrast is less salient and distinctive than nasalization.

In view of the close correspondence between putative phonological constraints and phonetic processes, one might rightly wonder whether the very distinction between the phonological and phonetic levels is justified (Flemming, 2001). Indeed, arguments for the distinction typically cite the contrast between “categorical” and “gradient” representations (for phonology and phonetics, respectively). But as noted earlier, many aspects of phonological knowledge, are, in fact, gradient, whereas binary contrasts could well emerge from gradient continuous distinctions (Maye et al., 2002; McClelland, 2009). In short, a phonetic account of the sound pattern of language would seem to lose little in observational adequacy, and gain much in explanatory power.

But while phonetic grounding offers important insights into the design of phonological systems, these observations do not demonstrate that the phonetic and phonological systems are one and the same. First, correlations (between the phonetic and phonological systems) should be kept distinct from causation – the possibility that phonological processes are molded by phonetic principles operating on-line. And indeed, the correlations in synchrony might have their source in diachrony, rooted in either natural or cultural evolution. Considering cultural evolution, phonetic factors could shape phonological systems by constraining the transmission of sound patterns across generations of speakers and hearers (Blevins, 2004). Phonetically senseless patterns will become extinct because they cannot be faithfully transmitted from speakers to hearers. In addition, phonetic preferences could also shape the design of the language system in phylogeny (e.g., as a result of natural selection) to yield universal grammatical constraints that are phonetically grounded (Berent, 2013a). While the mechanisms of human language evolution are presently unknown, the existence of such evolutionary pressures would be fully in line with many other cases of species-specific vocal patterns that likewise adapt to their sensorimotor channels (e.g., Suthers and Zollinger, 2004). Accordingly, the correlation between phonetic and phonological structure does not necessarily indicate an active casual role of phonetic in the synchronic grammar.

In fact, there are specific computational reasons to doubt this possibility. Phonological systems, as noted above, routinely encode functions such as identity, reduplication and correspondence – relations defined as operations over abstract variables. Moreover, people freely extend these functions across the board. An account of such generalizations requires mechanisms that ignore distinctions between class members (e.g., between members of the “consonantal” class), and that operate by combining discrete symbols in a manner that is structure sensitive (Fodor and Pylyshyn, 1988; Pinker and Prince, 1988). The phonetic system, by contrast, is an analog blending system that is exquisitely exemplar-based (Abler, 1989). In such systems, interaction is captured not by the combinatorial operations that are structure sensitive, but rather by blending, evident in tradeoff phenomena (Liberman et al., 1967). It is precisely the acute sensitivity of such system to continuous acoustic cues that explains why a phonetic account of phonology must encode rich phonetic structure.

Viewed in this fashion, the contrast between phonology and phonetics is not merely one between graded and categorical representations. Rather, it is a contrast between algebraic operations that are combinatorial and structure sensitive and analog systems that operate by blending. Whether these conflicting demands can be captured by a single computational system remains unknown. Paul Smolensky and colleagues (Smolensky et al., 2014), for example, explicitly address the challenge by integrating the Optimality-Theoretic grammar with computations that are sensitive to gradient phonetic detail. However, it is unclear whether this system can generalize relations (e.g., identity) as discussed here. In view of the sharp disparities between the computational characteristics of these two systems, a single system would be hardly trivial.

Beyond the computational challenge of integrating algebraic and analog systems, there are also substantive reasons to question this unification. One line of evidence comes from findings suggesting that phonology might occasionally betray phonetic pressures (de Lacy, 2006b; de Lacy and Kingston, 2013) – such an outcome counters a unified phonology/phonetic account. A second challenge is presented by the potential for amodal phonological universals. At the heart of the phonetic alternative is the hypothesis that the properties of phonological systems can be captured by the phonetic restrictions on the sound patterns of language. Yet, phonology is not invariably transmitted by sound. Every established sign language exhibits a manual sign system (Stokoe, 1960). But despite the sharp contrast in modality, signed and spoken phonologies share primitives and constraints (e.g., Sandler and Lillo-Martin, 2006; Berent et al., 2013a). The existence of putative phonological universals in the face of phonetic disparities suggests that the phonological and phonetic systems are distinct. Additional such dissociations are presented by neurological disorders. Dyslexia, for instance, is known to impair the phonetic system, but recent evidence suggests that the phonological system might be spared (Berent et al., 2012a; Berent et al., 2013b).

The dissociations between the phonological system and its phonetic speech channel, on the one hand, and the putative associations between signed and spoken phonology, on the other, present formidable challenges to single-system accounts. These observations, however, all naturally fall out from the hypothesis that phonological systems are algebraic, abstract and distinct from the phonetic system.

21.6.3 A dual-route account of phonology

The discussion in the previous sections contrasted an algebraic, structure-sensitive account of phonology with alternatives that are exemplar-based – either associative or phonetically grounded. In defense of the algebraic hypothesis, we have suggested that algebraic mechanisms are necessary to capture the scope of phonological generalizations, and that such generalizations cannot be reduced to the statistical structure of the lexicon. The sharp contrast between the computational properties of phonology and phonetic systems, and their multiple dissociations likewise challenge the attempt to eliminate algebraic mechanisms in favor of phonetic-based explanations.

But while algebraic mechanisms appear necessary to capture phonological knowledge, they are clearly not sufficient. As shown in earlier sections, some aspects of phonological knowledge are clearly exemplar-based, and those facts ought to be captured by phonological theory. So rather than dispensing with the grammar, such facts call for a broad integrative perspective that includes multiple computational mechanisms. The results presented in this chapter suggest that this rich phonological system must include an algebraic grammatical system at its core.

References

Abler, W. L. 1989. On the particulate principle of self-diversifying systems. *Journal of Social and Biological Systems*, 12, 1–13.
Adriaans, F. & R. Kager. 2010. Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory & Language*, 62, 311–331.
Albright, A. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology*, 26, 9–41.
Alderete, J., J. Beckman, L. Benua, A. Gnanadesikan, J. McCarthy & S. Urbanczyk. 1999. Reduplication with fixed segmentism. *Linguistic Inquiry*, 30, 327–364.
Archangeli, D. & D. Pulleyblank. 1994. *Grounded phonology*, Cambridge, MA, MIT Press.
Balota, D. & J. Chumbly. 1984. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 340–357.
Bat-El, O. 2006. Consonant identity and consonant copy: The segmental and prosodic structure of Hebrew reduplication. *Linguistic Inquiry*, 37, 179–210.
Berent, I. 2002. Identity avoidance in the Hebrew lexicon: Implications for symbolic accounts of word formation. *Brain and Language*, 81, 326–341.
Berent, I. 2013a. *The phonological mind*, Cambridge, Cambridge University Press.
Berent, I. 2013b. The phonological mind. *Trends In Cognitive Sciences*, 17, 319–327.
Berent, I., U. Bibi & J. Tzelgov. 2000. The autonomous computation of linguistic structure in reading: Evidence from the Stroop task. A paper presented at the 41st meeting of the Psychonomic Society, New Orleans: LA.
Berent, I., U. Bibi & J. Tzelgov. 2006. The autonomous computation of linguistic structure in reading: Evidence from the Stroop task. *The Mental Lexicon*, 1, 201–230.
Berent, I., A. Dupuis & D. Brentari. 2013a. Amodal aspects of linguistic design. *PLoS ONE*, 8, e60617.
Berent, I., D. L. Everett & J. Shimron. 2001a. Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology*, 42, 1–60.
Berent, I., G. F. Marcus, J. Shimron & A. I. Gafos. 2002. The scope of linguistic generalizations: Evidence from Hebrew word formation. *Cognition*, 83, 113–39.
Berent, I. & J. Shimron. 1997. The representation of Hebrew words: Evidence from the Obligatory Contour Principle. *Cognition*, 64, 39–72.
Berent, I. & J. Shimron. 2003. Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics*, 39, 31–55.
Berent, I., J. Shimron & V. Vaknin. 2001b. Phonological constraints on reading: Evidence from the Obligatory Contour Principle. *Journal of Memory and Language*, 44, 644–665.
Berent, I., V. Vaknin & G. Marcus. 2007. Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition*, 104, 254–286.
Berent, I., V. Vaknin & J. Shimron. 2004. Does a theory of language need a grammar? Evidence from Hebrew root structure. *Brain and Language*, 90, 170–182.
Berent, I., V. Vaknin-Nusbaum, E. Balaban & A. Galaburda. 2012a. Dyslexia impairs speech recognition but can spare phonological competence. *PLoS ONE*, 7, e44875.
Berent, I., V. Vaknin-Nusbaum, E. Balaban & A. Galaburda. 2013b. Phonological generalizations in dyslexia: The phonological grammar may not be impaired. *Cognitive Neuropsychology*, 30, 285–310.
Berent, I., C. Wilson, G. Marcus & D. Bemis. 2012b. On the role of variables in phonology: Remarks on Hayes and Wilson. *Linguistic Inquiry*, 43, 97–119.
Blevins, J. 2004. *Evolutionary phonology*, Cambridge, Cambridge University Press.
Browman, C. P. & L. Goldstein. 1989. Articulatory gestures as phonological units. *Phonology*, 6, 201–251.
Bybee, J. & J. L. McClelland. 2005. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Linguistic Review*, 22, 381–410.

Chomsky, N. 1972. *Language and mind*, New York, Harcourt Brace Jovanovich.
Chomsky, N. & M. Halle. 1968. *The sound pattern of English*, New York, Harper & Row.
Cole, J. 2009. Emergent feature structures: Harmony systems in exemplar models of phonology. *Language Sciences*, 31, 144–160.
Coleman, J. & J. Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. *In:* J. Coleman (ed) *Third meeting of the ACL special interest group in computational phonology: Proceedings of the workshop* (pp. 49–56), East Stroudsburg, PA, Association for computational linguistics.
De Lacy, P. 2006a. *Markedness: Reduction and preservation in phonology*, Cambridge, New York, Cambridge University Press.
De Lacy, P. 2006b. Transmissibility and the role of the phonological component. *Theoretical Linguistics*, 32, 185–196.
De Lacy, P. & J. Kingston. 2013. Synchronic explanation. *Natural Language and Linguistic Theory*, 31, 287–335.
Doignon, N. & D. Zagar. 2005. Illusory conjunctions in French: The nature of sublexical units in visual word recognition. *Language and Cognitive Processes*, 20, 443–464.
Flemming, E. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, 18, 7–44.
Flemming, E. 2004. Contrast and perceptual distintiveness. *In:* B. Hayes, R. M. Kirchner & D. Steriade (eds) *Phonetically based phonology* (pp. 232–276), Cambridge, Cambridge University Press.
Fodor, J. A. 1975. *The language of thought*, Cambridge, MA, Harvard University Press.
Fodor, J. A. & Z. Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–71.
Frisch, S. A., J. B. Pierrehumbert & M. B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, 22, 197–228.
Gafos, A. I. 1999. *The articulatory basis of locality in phonology*, New York, Garland Publishers.
Gallagher, G. 2013. Learning the identity effect as an artificial language: Bias and generalisation. *Phonology*, 30, 253–295.
Gervain, J., I. Berent & J. Werker. 2012. Binding at birth: Newborns detect identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24, 564–574.
Hahn, U. & T. M. Bailey. 2005. What makes words sound similar? *Cognition*, 97, 227–267.
Harm, M. W. & M. S. Seidenberg. 1999. Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106, 491–528.
Hayes, B., R. M. Kirchner & D. Steriade (eds). 2004. *Phonetically based phonology*, Cambridge, Cambridge University Press.
Hayes, B. & C. Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39, 379–440.
Hayes, B. P. 1999. Phonetically driven phonology: The role of Optimality Theory and inductive grounding. *In:* E. A. M. Danell, F. Newmeyer, M. Noonan & K. W. Wheatley (eds) *Formalism and functionalism in linguistics* (pp. 243–285), Amsterdam, Benjamins.
Inkelas, S. 2008. The dual theory of reduplication. *Linguistics*, 46, 351–401.
Joanisse, M. F. & M. S. Seidenberg. 2003. Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain Lang*, 86, 40–56.
Jun, J. 2004. Place assimilation. *In:* B. Hayes, R. Kirchner & D. Steriade (eds) *Phonetically based phonology* (pp. 58–86), Cambridge, Cambridge University Press.
Kim, J. J., G. F. Marcus, S. Pinker, M. Hollander & M. Coppola. 1994. Sensitivity of children's inflection to grammatical structure. *Journal of Child Language*, 21, 173–209.
Kucera, H. & W. Francis. 1967. *Computational analysis of present-day American English*, Providence, RI, Brown University Press.
Leben, W. 1973. *Suprasegmental phonology*, Cambridge, MA, MIT Press.
Liberman, A. M., F. S. Cooper, D. P. Shankweiler & M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review*, 74, 431–461.

Lombardi, L. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language & Linguistic Theory*, 17, 267–302.

Luce, P. A. & D. B. Pisoni. 1998. Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19, 1–36.

Marcus, G. F. 1998. Rethinking eliminative connectionism. *Cognitive Psychology*, 37, 243–282.

Marcus, G. F. 2001. *The algebraic mind: Integrating connectionism and cognitive science*, Cambridge, MIT Press.

Marcus, G. F., U. Brinkmann, H. Clahsen, R. Wiese & S. Pinker. 1995. German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189–256.

Marcus, G. F., S. Vijayan, S. Bandi Rao & P. M. Vishton. 1999. Rule learning by seven-month-old infants. *Science*, 283, 77–80.

Mattingly, I. G. 1981. Phonetic representation and speech synthesis by rule. *In:* T. Myers, J. Laver & J. Anderson (eds) *The cognitive representation of speech* (pp. 415–420), Amsterdam, North Holland.

Maye, J., J. F. Werker & L. Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.

McCarthy, J. J. 1979. *Formal problems in Semitic phonology and mophology. Doctoral dissertation, MIT*., New York, 1985, Garland Press.

McCarthy, J. J. 1994. The phonetics and phonology of Semitic pharyngeals. *In:* P. Keating (ed) *Papers in Laboratory Phonology III* (pp. 191–233), Cambridge, Cambridge University Press.

McClelland, J. L. 2009. Phonology and perception: A cognitive scientist's perspective. *In:* P. Boersma & S. Hamann (eds) *Phonology in perception* (pp. 293–314), Berlin, Mouton De Gruyter.

McClelland, J. L. & K. Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, 6, 465–472.

McCutchen, D., L. C. Bell, I. M. France & C. A. Perfetti. 1991. Phoneme-specific interference in reading: The tongue-twister effect revisited. *Reading Research Quarterly*, 26, 87–103.

Monaghan, P. & R. C. Shillcock. 2003. Connectionist modelling of the separable processing of consonants and vowels. *Brain Lang*, 86, 83–98.

Ohala, J. J. 1975. Phonetic explanations for nasal sound patterns. *In:* C. A. Ferguson, L. M. Hyman & J. J. Ohala (eds) *Nasalfest: Papers from a symposium on nasals and nasalization* (pp. 289–316), Stanford, Language Universals Project.

Perea, M. & M. Carreiras. 1998. Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 134–144.

Perfetti, C. A. & L. Bell. 1991. Phonemic activation during the first 40 ms. of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, 30, 473–485.

Perfetti, C. A., L. C. Bell & S. M. Delaney. 1988. Automatic (pre-lexical) phonetic activation in silent reading: Evidence from backward masking. *Journal of Memory and Language*, 32, 57–68.

Pierrehumbert, J. B. 2001. Stochastic phonology. *GLOT*, 5–9, 1–13.

Pinker, S. 1991. Rules of language. *Science*, 253, 530–535.

Pinker, S. 1999. *Words and rules: The ingredients of language*, New York, Basic Books.

Pinker, S. & A. Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73–193.

Prasada, S. & S. Pinker. 1993. Generalization of regular and irregular morphological patters. *Language and Cognitive Processes*, 8, 1–55.

Prince, A. & P. Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*, Malden, MA, Blackwell Pub.

Prince, A. & P. Smolensky. 1997. Optimality: From neural networks to universal grammar. *Science*, 275, 1604–1610.

Pylyshyn, Z. 1984. *Computation and cognition: Towards a foundation for cognitive science*, Cambridge, MIT Press.

Rapp, B. 1992. The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language*, 31, 33–53.

Rose, S. & R. Walker. 2004. A typology of consonant agreement as correspondence. *Language*, 80, 475–531.

Rumelhart, D. E. & J. McClelland. 1986. On learning past tense of English verbs: Implicit rules or parallel distributed processing? *In:* D. E. Rumelhart, J. McClelland & T. P. R. Group (eds) *Parallel distributed processing: Explorations in the microstructure of cognition* (pp. 216–271), Cambridge, MA, MIT Press.

Sandler, W. & D. C. Lillo-Martin. 2006. *Sign language and linguistic universals*, Cambridge, Cambridge University Press.

Seidenberg, M. 1987. Sublexical structures in visual word recognition: Access units of orthographic redundancy? *In:* M. Coltheart (ed) *Attention and performance XII: The psychology of reading* (pp. 245–263), Hillsdale, NJ, Erlbaum.

Shibatani, M. 1973. The role of surface phonetic constraints in generative phonology. *Language*, 49, 87–106.

Smolensky, P., M. Goldrick & D. Mathis. 2014. Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science*, 38, 1102–1138.

Smolensky, P. & G. Legendre. 2006. *The harmonic mind: From neural computation to Optimality-theoretic grammar*, Cambridge, MA, MIT Press.

Solso, R. L. & C. L. Juel. 1980. Positional frequency and versatility of bigrams for two- through nine-letter English words. *Behavior Research Methods & Instrumentation*, 12, 297–343.

Steriade, D. 1997. *Phonetics in phonology: The case of laryngeal neutralization*, Unpublished manuscript.

Steriade, D. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization.

Stokoe, W. C., Jr. 1960. Sign language structure: An outline of the visual communication systems of the American Deaf. *Journal of Deaf Studies and Deaf Education*, 10, 3–37.

Suthers, R. A. & S. A. Zollinger. 2004. Producing song: The vocal apparatus. *Annals of the New York Academy of Sciences*, 1016, 109–129.

Suzuki, K. 1998. *A typological investigation of dissimilation. Doctoral dissertation*, University of Arizona.

Taft, M. 1979. Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7, 263–272.

Turing, A. M. 1936. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 24, 230–265.

Turing, A. M. 1950. Computing machinery and intelligence. *Mind: A Quarterly Review of Philosophy*, 59, 433–460.

Vitevitch, M. S., P. A. Luce, D. B. Pisoni & E. T. Auer. 1999. Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain & Language*, 68, 306–311.

Wright, R. 2004. A review of perceptual cues and cue robustness. *In:* B. Hayes, R. M. Kirchner & D. Steriade (eds) *Phonetically based phonology*, Cambridge: Cambridge University Press.

Yip, M. 1988. The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry*, 19, 65–100.

Zhang, J. 2004. The role of contrast-specific and language-specific phonetics in contour tone distribution. *In:* B. Hayes, R. Kirchner & D. Steriade (eds) *Phonetically based phonology* (pp. 157–190), Cambridge, Cambridge University Press.

Ziegler, J. C., M. Muneaux & J. Grainger. 2003. Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, 48, 779–793.

22
Statistical phonology[1]

Michael Hammond

22.1 Overview

Classical phonological theory was founded on the generative distinction between competence and performance (Chomsky & Halle 1968). Phonological "facts" were intuitions of grammaticality expressed by native speakers. There are very good reasons to adopt some version of this view; all the classical arguments advanced by, e.g. Chomsky (1957) or Chomsky (1965), apply just as well to phonology as to syntax. The set of phonologically grammatical words or utterances in a language is infinite and thus cannot be extracted from any imaginable finite corpus. In addition, the existence of speech errors means that the set of occurring words or utterances is not a proper subset of the set of grammatical words or utterances.

However, as in syntax, this focus on grammaticality judgments entails no real role for quantitative methods. The theoretical machinery of the day thus had no quantitative element. In addition, data gathering techniques of the time, e.g. collecting grammatical judgments, were also not quantitative in nature.

This all changed very quickly when phonologists started looking at data beyond grammaticality judgments. The first data to challenge orthodox assumptions were sociolinguistic, where it was argued that traditional generative data gathering techniques could not be used because of the effects of dialect contact (Labov 1969), essentially the *observer's paradox*. Other data that seem to cry out for a quantitative treatment include corpus data and experimental data from phonetics or psycholinguistics.

One could argue – and a number of researchers have and still do – that apparent quantitative effects in phonology are outside the grammar, and are ultimately performance-based. If we continue to maintain a distinction between competence and performance, then surely at least some apparent quantitative effects in phonology are more properly seen as part of performance. There are, however, a number of reasons to believe that at least some quantitative effects should be a part of phonological competence, whether or not one takes the position that the competence–performance distinction should be abandoned.

First, there are statistical effects that mirror categorical effects. For example, categorical markedness effects in one language are mirrored by statistical facts in another. For example, voiced stops are generally considered marked with respect to voiceless stops. In categorical terms, this expresses itself in the existence of languages with voiceless stops and no voiced stops, but not the contrary (Greenberg 1978). In statistical terms, languages that have both stop series will typically have more voiceless stops than voiced stops.

Second, it's well established now that statistical regularities affect acquisition. Thus, for example, more frequent sounds are learned earlier than less frequent sounds (Zamuner 2001; Zamuner et al. 2004). Similarly, statistical distribution over a perceptual continuum will affect where phoneme boundaries are drawn (Maye 2000), e.g. if items exhibit a bimodal distribution, subjects are more likely to judge that they belong to two categories than if they exhibit a monomodal distribution.

Third, statistical distributions affect grammaticality judgments. The well-formedness of an item is a function of the statistical frequency of its phonological components (Greenberg & Jenkins 1964; Ohala & Ohala 1986; Frisch et al. 2000).

In this paper, we review the various tacks that have been taken to incorporate statistical regularities in phonological theory. We begin with the classic variable rule literature in rule-based phonology. We then turn to statistical approaches in constraint-based theories. These run the gamut from counting constraint violations to variable constraint ranking to variable weights for constraints and variable constraint sets. We also treat various proposals for treating statistical phonotactics and lexical frequency. We conclude with a discussion of where this may lead.

22.2 Variable rules

Labov (1969) argues that sociolinguistic data cannot be collected using standard grammaticality judgments. The problem, he maintains, is that those judgments would be influenced by the dominant dialect. Sociolinguistic data must then be collected with interviews and observation, and these necessarily provide quantitative data. With a grammaticality judgment, a construction is or is not grammatical. With some construction observed in context, we simply have a count for how often it occurs in various contexts.

To accommodate this kind of quantitative data, Labov develops a theory of *variable rules*. At the simplest level, every rule is associated with some value k_0 which governs how likely the rule is to apply generally (φ). As k_0 increases, likelihood decreases:

(1) $\varphi = 1 - k_0$

Imagine, for example, we have a rule like the following (Labov 1969: 738). Here, parentheses indicate that the rule is variable.

(2) $X \rightarrow (Y) / \begin{bmatrix} \alpha F_i \\ \vdots \end{bmatrix} \begin{bmatrix} \overline{\quad} \\ \gamma F_j \\ \vdots \end{bmatrix} \begin{bmatrix} \beta F_k \\ \vdots \\ \nu F_n \end{bmatrix}$

The Greek letters here range over {+, −}. Each feature of the environment is also associated with a value $k_1, k_2, \ldots k_n$, so that:

(3) $\varphi = 1 - (k_0 - \alpha k_1 - \beta k_2 \ldots \nu k_n)$

On this view, the likelihood of a rule applying is an inverse function of the overall likelihood of that rule (k_0) plus or minus the contributions of each of the features in the environment of the rule. For some feature k_i with value α, we define the fraction of the set of sentences for which $\alpha = +$ as $\varphi(\alpha)$. The complementary set is $\varphi(\sim\alpha)$.

Let's look at a *very* simple hypothetical example to see how this works. We simplify (2) as follows.

(4) $X \rightarrow (Y) / [\alpha F]\_\_$

Assume in addition that the constant for F is k_1, $k_0 = .3$, and $k_1 = .2$. If, in some environment, we have $+F$, then we have:

(5)
$$\begin{aligned} \varphi &= 1 - (.3 - (+.2)) \\ &= (1 - (.3 - .2) \\ &= 1 - .1 \\ &= .9 \end{aligned}$$

If, in some environment, we have $-F$, then we have:

(6)
$$\begin{aligned} \varphi &= 1 - (.3 - (-.2)) \\ &= 1 - (.3 + .2) \\ &= 1 - .5 \\ &= .5 \end{aligned}$$

That is, rule (4) is more likely to apply when $+F$ occurs to the left.

Labov uses a special notation, an asterisk, to indicate that a rule is obligatory in some context, that the presence of some feature overrides whatever variability might be introduced by other features. In the following example, vowel syncope is obligatory before nasals and after pronouns.

(7)
$$\text{ə} \rightarrow (\varphi) / [\text{*pro}]\#\# \begin{bmatrix} \overline{+\text{T}} \end{bmatrix} \begin{matrix} \text{C}_0^1 \\ [\text{*nas}] \end{matrix} \#\# \begin{bmatrix} \alpha\text{Vb} \\ \beta\text{gn} \end{bmatrix}$$

This then gives the following interpretations for the coefficients that a feature can bear in a rule:

(8)

Rule contains notation . . .	$\varphi(F_1)$	$\varphi(\sim F_1)$
$+F_1$	φ	0
$-F_1$	0	φ
αF_1	$1-(k_0-k_1\ldots)$	$1-(k_0+k_1\ldots)$
$-\alpha F_1$	$1-(k_0+k_1\ldots)$	$1-(k_0-k_1\ldots)$
$*F_1$	1	φ
$-*F_1$	φ	1

An explicit plus or minus means that the feature must be present or absent respectively. If that condition holds, then other values govern the precise applicability of the rule. If not, the rule simply cannot apply. Greek letters indicate the calculations we've already discussed. An asterisk indicates that the rule is categorical if the relevant value is present.

Returning to our simple example in (4), let us revise that rule as follows:

(9) $X \rightarrow (Y) / \begin{bmatrix} \alpha F \\ {*G} \end{bmatrix}$ __

There are four cases to consider:

(10)	$+F$	$+G$	1
	$+F$	$-G$	.9
	$-F$	$+G$	1
	$-F$	$-G$	.5

If the value for G is $+G$, then the rule must apply. If the value is $-G$, then we get the same calculations as before.

Finally, the value of the constants for each Greek variable are ordered:

(11) $k_1 > k_2 > \ldots > k_n$

More specifically:

(12) Postulate of Geometric Ordering
If $\chi_1, \chi_2, \ldots, \chi_n$ are variable constraints upon a rule r, then for any given $\chi_1, \chi_2, \ldots, \chi_{i-1}, \varphi_r(\chi_i) > \varphi(\sim \chi_i)$.

In other words, each constraint in the hierarchy outweighs the effects of all constraints below it. One set of values that satisfies this requirement for three constraints is: $k_0 = \frac{1}{2}(.5)$, $k_1 = \frac{1}{4}(.25)$, $k_2 = \frac{1}{8}(.125)$, $k_3 = \frac{1}{16}(.0625)$.

(13)

Constraints	Overall	Sum
$+\alpha, +\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}$	.94
$+\alpha, +\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}-\frac{1}{16}$	.81
$+\alpha, -\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{1}{16}$	.69
$+\alpha, -\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}-\frac{1}{16}$	.56
$-\alpha, +\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}+\frac{1}{16}$	.44
$-\alpha, +\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{1}{16}$	.31
$-\alpha, -\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}+\frac{1}{16}$	.19
$-\alpha, -\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\frac{1}{16}$	.06

Since the feature value coefficients are added together, there is a risk that the result will not be a legal probability value, that $\varphi < 0$ or $\varphi > 1$. Geometric ordering avoids this problem. Consider, for example, what would happen if we had these values instead: $k_0 = \frac{1}{2}$ (.5), $k_1 = \frac{1}{4}$ (.25), $k_2 = \frac{1}{8}$ (.125), $k_3 = \frac{3}{4}$ (.75).

(14)

Constraints	Overall	Sum
$+\alpha, +\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{3}{4}$	1.65
$+\alpha, +\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}-\frac{3}{4}$	0.125
$+\alpha, -\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{3}{4}$	1.375
$+\alpha, -\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}-\frac{3}{4}$	−0.126
$-\alpha, +\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}+\frac{3}{4}$	1.125
$-\alpha, +\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{3}{4}$	−0.375
$-\alpha, -\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}+\frac{3}{4}$	0.875
$-\alpha, -\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\frac{3}{4}$	−0.625

This results in values that cannot be probability values.

Cedergren & Sankoff (1974) treat variable rules from a mathematical perspective. They start from the observation that "Labov has . . . discovered the highly significant generalization that the presence of a given feature or subcategory tends to affect rule frequency in a probabilistically uniform way in all the environments containing it" (p. 336).

There are a number of ways the model can be expressed mathematically. The simplest is the model where the contribution of all features or subcategories are summed:

(15) $$p = p_0 + \alpha_i + \alpha_j + \ldots$$

This is the same structure as ANOVA, and it's what Labov (1969) used. As we've seen in (14), additive models can produce nonsensical probability values below zero or greater than 1. Strong geometric ordering is one way to avoid this. Another is to truncate values that fall outside the range of interpretability. Values outside the range of legal probability values would be replaced with legal (extreme) values. Taking the values in (14), we would get the following:

(16)

Constraints	Overall	Sum	Truncated
$+\alpha, +\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{3}{4}$	1.65	1
$+\alpha, +\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}-\frac{3}{4}$	0.125	0.125
$+\alpha, -\beta, +\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{3}{4}$	1.375	1
$+\alpha, -\beta, -\gamma$	$\frac{1}{2}+\frac{1}{4}-\frac{1}{8}-\frac{3}{4}$	−0.126	0
$-\alpha, +\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}+\frac{3}{4}$	1.125	1
$-\alpha, +\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{3}{4}$	−0.375	0
$-\alpha, -\beta, +\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}+\frac{3}{4}$	0.875	0.875
$-\alpha, -\beta, -\gamma$	$\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\frac{3}{4}$	−0.625	0

Cedergren & Sankoff propose alternative multiplicative models which also address this problem: here we take the product of the values for each feature. Since the terms are all legal probability values, only a legal probability value can result.

(17) $(1 - p) = (1 - p_0) \times (1 - p_i) \times (1 - p_j) \times \ldots$

Cedergren & Sankoff consider various modifications of this. One dimension they explore is whether to model the rule's application or its non-application. They also consider an additive logarithmic model:

(18) $\log p = \beta_0 + \beta_a + \beta_b + \ldots + \beta_n$

Sankoff & Labov (1979) develop the statistical model further as well.

So far, we've just considered the general formal and mathematical structure of the model; let's now consider how it's used. Guy (1991) is an extremely impressive application of this approach. He combines variable rules with Lexical Phonology (Kiparsky 1982) to capture the distribution of t/d-deletion in different morphological categories. The facts are as follows. Word-final t/d are deleted in monomorphemic forms at a higher rate than in morphologically irregular past tense forms. The latter exhibit a higher rate of deletion than regular past tense forms.

(19)

Class	N Total	N Deleted	% Deleted	% Retained
Monomorphemic (e.g. *mist, pact*)	658	251	38.1	61.9
Past semiweak (e.g. *left, told*)	56	19	33.9	66.1
Past regular (e.g. *missed, packed*)	181	29	16.0	84.0

The theory of Lexical Phonology holds that the phonology and morphology are organized into levels or strata with different morphological processes associated with each stratum. In English, irregular past tenses are created at an earlier level than regular past tenses. The rule of t/d-deletion is associated with all levels of the English lexicon. This means it gets multiple chances to apply. (For this to work, one has to set aside notions of Strict Cyclicity.) What this means is that monomorphemic forms get at least three chances to undergo deletion, irregular past tense forms get two, and regular past tense forms get only one chance.

Let's represent the probability of a rule applying as p_a. If that rule applies at some level, it will leave a remainder: $1 - p_a$. After n levels, the proportion of forms to which the rule hasn't applied will be $(1 - p_a)^n$. For example, if the probability of the rule is .4, then after one application .4 of the forms will have undergone it and $1 - .4 = .6$ will not have. After two applications, $(1 - .4)^2 = .36$ will not have undergone it, after three: $(1 - .4)^3 = .216$, after four: $(1 - .4)^4 = .1296$, etc. This makes the wonderful prediction that the rate of application for later strata/levels should directly reflect the rate of application for the first stratum.

Let's assume this is the rule:

(20) $\{t,d\} \rightarrow \langle \phi \rangle / \text{C\_\_}]$

The residue for regular past tense forms is 84%. This, in turn, predicts that the rate of retention for irregular forms is $.84^2 = .7056$, and the rate of retention for monomorphemic forms is $.84^3 = .5927$. These values are quite close to the attested values in (19) of 66.1 and 61.9 respectively, and Guy takes this as evidence that his hypothesis is correct. (He goes on to test the model using different statistical tests.)

22.3 Constraints

While variable rules enjoyed and continue to enjoy wide use in the sociolinguistic literature, they were never really integrated into mainstream phonology. As mainstream generative phonology moved from traditional linear rule-based phonology to non-linear autosegmental and metrical approaches, and then to constraint-based approaches, a linear rule-based approach to variability fit that mainstream less and less. However, there has been significant subsequent attention devoted to variation in constraint-based terms. (See Coetzee & Pater 2011 for an excellent review.) We therefore need to understand these in some detail as these approaches involve specific manipulations of the mathematics behind constraints.

An early paper that makes a compelling case for constraints is Kisseberth (1970). Kisseberth demonstrates that there are a number of phonological rules in Yawelmani that conspire together to achieve specific configurations of consonants and vowels (syllable structure).

Using the technology of the day, these rules could not be expressed as a single rule, hence missing the generalization that they all have the same general consequence.

For example, there is a rule that inserts a vowel in a word-final two-consonant cluster or in any three-consonant cluster. There is also a rule that deletes a consonant when it occurs in a heteromorphemic three-consonant cluster.

(21) $\phi \rightarrow V / C \_\_ C \left\{ \begin{smallmatrix} \# \\ C \end{smallmatrix} \right\}$

(22) $C \rightarrow \phi / C \left\{ \begin{smallmatrix} C+\_\_ \\ +\_\_C \end{smallmatrix} \right\}$

Both rules have the effect of avoiding three-consonant clusters, but this cannot be expressed using the rule formalism of the time. We might think of the rules – and the languages – as being subject to a more global general constraint against such clusters.

Ito (1989) argued for this position explicitly. Syllabic templates – which are independent from phonological rules – are proposed and these templates derive the kinds of effects Kisseberth observed years earlier. These templates operate as constraints on the phonological derivation.

In the framework of *Optimality Theory* (henceforth OT; Prince & Smolensky 1993; McCarthy & Prince 1993), the role of constraints is expanded so that there are no rules left, or only a single rule: *generate* (Gen). It is important to examine this framework more closely to understand the mathematics behind it as it is precisely in terms of the structure and mathematics of this theory that many modern statistical phonological theories can be described.

The basic logic of OT proceeds as follows. (This is a simplified version of the formalism presented in Prince & Smolensky 1993.) First, for any input or underlying form, we consider every possible output, or possible pronunciation. This pairing of possible outputs with an input form is performed by the function Gen. There is then a set of universal constraints that assign penalties or violations to the different pairings. As a first approximation, we say that the *harmony* of a candidate is inversely proportional to the number of violations it incurs.

For example, we might have the input /pak/ with a final consonant and the possible outputs [pa] and [pak] where that final consonant is either pronounced as a coda or deleted. A constraint against codas, NoCoda, would assign a violation to the input–output pairing /pak/-[pak].

(23)

	/pak/	NoCoda
☞	pa	
	pak	*!

Constraints can be violated more than once for any particular input–output pair. When we consider only one constraint *C*, a pair *A* is more harmonic than another pair *B* if *A* has fewer violations of *C* than *B*. We represent this as:

(24) $A \succ B \text{ iff } |A|_C < |B|_C$

In the example above, representing pairs just by the candidate form, we have:

(25) $\text{pa} \succ \text{pak because } |\text{pa}|_{\textsc{NoCoda}} < |\text{pak}|_{\textsc{NoCoda}}$

If we had a constraint that could exhibit multiple violations, winning status would be determined by the same procedure. Consider the following schematic tableau:

(26)

	I(nput)	C(onstraint)
☞	O(utput)$_1$	*
	O(utput)$_2$	**!

Here O_1 wins:

(27) $O_1 \succ O_2$ because $|O_1|_C < |O_2|_C$

The definition of harmony becomes more complex when additional constraints are involved. Let's add an additional candidate to the mix: [pakə], where the final consonant is saved from codahood by epenthesis. Let's assume, moreover, that it is this candidate that wins. Several constraint rankings are consistent with this result, but let's adopt the following one:

(28)

	/pak/	NoCoda	Max	Dep
	pa		*!	
	pak	*!		
☞	pakə			*

Here, [pakə] wins because the other candidates have more violations of higher-ranked constraints.

Following Prince & Smolensky in general terms, each input–output pair is associated with a vector of constraint violation counts ($n_1, n_2, \ldots, n_n$). For example, in the tableau above, [pa] has the vector (0, 1, 0). A candidate A is more harmonic than a candidate B iff for some point in their two vectors $n_j^A < n_j^B$ and for all values k in $(1, \ldots, nj - 1)$, we have $n_k^A = n_k^B$.

(29)

a. $n_1^{\text{pa}} < n_1^{\text{pak}}$ pa ≻ pak

b. $n_1^{\text{pakə}} < n_1^{\text{pak}}$ pakə ≻ pak

c. $n_2^{\text{pakə}} < n_2^{\text{pa}}$ pakə ≻ pa
$n_1^{\text{pakə}} = n_1^{\text{pa}}$

Since [pakə] has the highest harmony by these calculations, it is the winning candidate.

There's an alternative formalization that allows us to dispense with vectors and anticipates developments to come, so let's take a look at it. Given a constraint system with a finite sequence of ranked constraints ($C_1, C_2, \ldots, C_n$), we associate weights with each constraint ($w_1, w_2, \ldots, w_n$). The harmony value $\mathcal{H}$ for any input–output pair is the sum of the products of the respective weights and violations:

(30) $\mathcal{H} = n_1 w_1 + n_2 w_2 + \ldots + n_n w_n$

A candidate A is more harmonic than a candidate B iff $\mathcal{H}_{\text{A}} < \mathcal{H}_{\text{B}}$.

A key property must hold for this to effectively capture *strict* ranking – where there are no "trade-offs" – which is that the weight w_i for any constraint must exceed $m_j\, w_j$, where m_j

is the maximum number of violations possible for the next constraint C_j in the ranking and w_j is the weight of that constraint. The last/lowest constraint in the ranking is not subject to this constraint and can be any positive value. (Note that an infinite number of violations may be possible for some constraint, so the weights may involve infinities of different sizes on this formalization.)

Returning to our simple example in (28), let's set the weights at (5, 3, 1). We calculate the harmony values as follows:

(31) pa $0 \times 5 + 1 \times 3 + 0 \times 1 = 3$
pak $1 \times 5 + 0 \times 3 + 0 \times 1 = 5$
pakə $0 \times 5 + 0 \times 3 + 1 \times 1 = 1$

We obtain the same result.

Thus far, the OT system as presented has no particular statistical basis or interpretation. Golston (1998) was the first to add an explicitly statistical interpretation to the model. (See Coetzee 2008 for a similar idea.) This paper treats Middle English poetic meter. As with the sociolinguistic data we considered above, we cannot rely on traditional generative grammaticality judgments for the well-formedness of poetic lines. Instead, attested lines are assumed to be grammatical. More specifically, Golston argues that the frequency of occurrence of a line in the corpus correlates with its harmony; line types that exhibit more violations of higher-ranked constraints occur less often.

Returning once again to our schematic example (28), if this system were governed by the distributional regularities Golston hypothesizes, we would expect more instances of [pakə] than [pa], and more instances of [pa] than [pak]. Golston doesn't treat this in statistical detail, but it's easy to imagine how to do so. Harmony values define a distribution which can be compared to the actual distribution. We can convert the values in (31) to a statistical distribution by dividing each candidate's value by the sum of all the values: (.33, .55, .11). We can then use a test like Chi-square (χ^2) to test this predicted distribution against the occurring distribution.

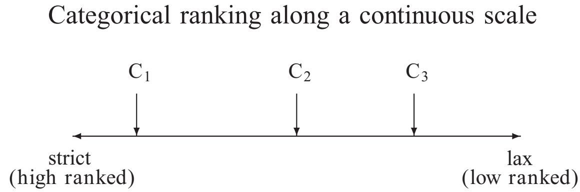
This, of course, is dependent on the weights that we have used. While these weights are limited by strict ranking, any finite set of strictly ranked constraints is consistent with an infinite set of possible weights. With this in mind, a more precise statistical test of Golston's proposal would first find the weights that fit the actual distribution best and then test those for goodness of fit. We will see models of explicitly this sort further on.

If, however, weights are subject to the constraint that the weight of some constraint C_i must be infinitely greater than the weight of the next constraint C_j in the ranking, then it's not clear that we'd be able to come up with a distribution that a test like χ^2 could be used with.

22.4 Stochastic OT

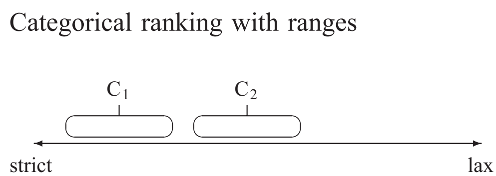
Boersma (1997) develops a very influential proposal for capturing ranking variation which involves manipulating the weights of constraints statistically. This is still a strict ranking system; the weights of different constraints must still be infinitely different. However, under certain circumstances, rankings can vary. Anticipating the proposal, here's how we might represent strict ranking. (As we've seen above, we must think of the weights for the constraints as all being infinitely different from each other.)

(32) Categorical ranking along a continuous scale



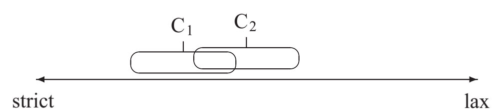
The idea is that constraints don't have a specific definite ranking, but a range of *possible* rankings. We can represent these by providing each constraint with a range of possible ranking values.

(33) Categorical ranking with ranges



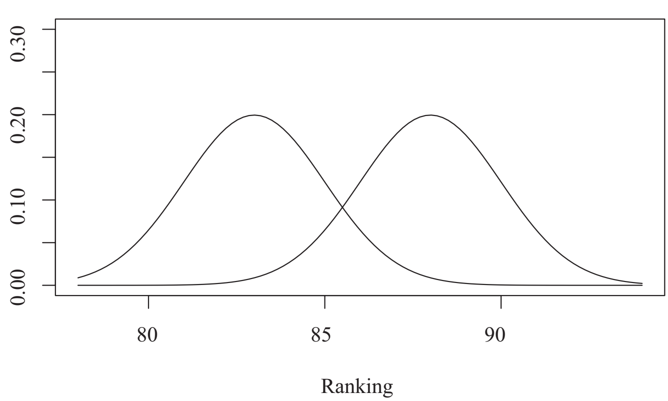
If those ranges overlap, then there is the possibility that in some percentage of the time the constraints can exhibit a different ranking with respect to each other. This is how variation is captured.

(34) Free ranking



In point of fact, Boersma proposes that the distribution of each constraint is normal. That is, for the range of ranking values that a constraint can exhibit, the constraint exhibits a normal distribution. We can represent this as in (35).

(35) Overlapping normal curves



On this view, the actual ranking of a constraint, its *disharmony* is given as:

(36) $dis = ranking + rs \times \mathrm{z}$

The *rs*, or *ranking spread*, variable is a constraint for the whole system, typically set to 2. The variable z is a Gaussian random variable with mean 0 and standard deviation 1. All constraints have the same mean and standard deviation.

The distribution of the harmony difference between two constraints c_1 and c_2 is:

(37) $dis_1 - dis_2 = r_1 - r_2 + rs \times (\mathrm{z}_1 - \mathrm{z}_2)$

That is, the actual relative ranking between two constraints is their ranking difference plus the difference between the values of their Gaussians times the *rs* term.

Generalizing over the joint distribution of the Gaussians, the probability that C_1 outranks C_2 is:

(38) $$\mathrm{p}(dis_1 > dis_2) = \frac{1}{2} \times \left(1 - \mathrm{erf}\left(\frac{1}{2}\sqrt{2} \times \frac{r_1 - r_2}{rs \times \sqrt{2}}\right)\right)$$

(Here, erf is the error function for a Gaussian.) We plot out these probabilities for different ranking differences as in (39).

(39)

$r_1 - r_2$	Probability
0	50%
1	36.2%
2	24%
3	14.4%
4	7.9%
5	3.9%
6	1.7%

An early non-statistical approach to variability within OT was "incomplete" rankings (Hammond 1994; Anttila 1995). In schematic terms, the basic idea was that the theory might specify some number of constraints $\{A, B, C\}$, but any particular grammar might only specify a partial ranking over those constraints, e.g. $A \gg C$ and $B \gg C$. Variability could occur where rankings were indeterminate, in this case, the ranking of *A* and *B*. One compelling argument for Stochastic OT over the incomplete rankings approach is that it provides a mechanism for describing cases where variation is not free, where the likelihood of a reranking of two constraints is not 50%. Another argument for Stochastic OT is that one can show that constraint rankings can be learned by relatively simple algorithms (Boersma & Hayes 2001).

Finally, the Stochastic OT proposal makes an extremely interesting prediction regarding the transitivity of constraint rankings. Under normal strict ranking, transitivity applies. That is, if constraint *A* outranks constraint *B*, and constraint *B* outranks constraint *C*, it follows that *A* outranks *C*. Stochastic OT allows us to extend this to the probabilistic domain. Imagine constraint *A* outranks constraint *B* such that the ranking difference is 1, and the probability of reranking is thus 36.2%. Imagine the same is true of constraints *B* and *C*. It then follows that the ranking difference between *A* and *C* is 2 and that those two constraints must have a 24% chance of reranking. (In very interesting unpublished work, Turton 2012 proposes a blend of Stochastic OT and Stratal OT which provides a more nuanced framework to treat effects like those discussed by Guy 1991 and reviewed in section 22.2 above.)

22.5 Harmonic OT

Pater (2009) and Potts et al. (2010) develop a version of OT where constraint weights are finite. (Similar ideas are developed in Legendre et al. 1990a, 1990b; Keller 2006; see also Goldsmith 1993.) The model in and of itself is not a model of variation and does not itself have a statistical component. However, it is a key step on the way to models that do.

The constraint-ranking logic is very similar to that of standard OT; each constraint C_k is associated with a weight w_k. Each candidate has a harmony value *H* calculated like this:

(40)
$$H = \sum_{k=1}^{K} s_k w_k$$

where s_k is the score for the candidate with respect to constraint C_k. The difference is that w_k can have a finite value.

This allows for two *trade-off* effects not possible in orthodox OT with its nonfinite constraint weights. First, multiple violations of a lower-ranked constraint can outweigh a single higher-ranked constraint. Imagine we have C_1 with $w_1 = 1$ and C_2 with $w_2 = .75$. With only a single violation of both constraints, we get the same effect as with strict ranking:

(41)

	Input	C_1	C_2
☞	$cand_1$		*
	$cand_2$	*!	

Here, $H_1 = (0 \times 1) + (1 \times .75) = .75$ while $H_2 = (1 \times 1) + (0 \times .75) = 1$.

On the other hand, when C_2 is violated twice, we get a different result:

(42)

	Input	C_1	C2
	$cand_1$		**!
☞	$cand_2$	*	

Here, $H_1 = (0 \times 1) + (2 \times .75) = 1.5$ while $H_2 = (1 \times 1) + (0 \times .75) = 1$.

We get the same kind of effect with multiple lower-ranked constraints. Imagine we have three constraints with weights $w_1 = 1$, $w_2 = .75$, $w_3 = .5$. A single violation of C_1 will outweigh a single violation of either C_2 or C_3.

(43)

	Input	C_1	C_2	C_3
☞	$cand_1$		*	
	$cand_2$	*!		

	Input	C_1	C_2	C_3
☞	$cand_1$			*
	$cand_2$	*!		

In the first case, we have $(0 \times 1) + (1 \times .75) + (0 \times .5) = .75$ vs. $(1 \times 1) + (0 \times .75) + (0 \times .5) = 1$. In the second case, we have $(0 \times 1) + (0 \times .75) + (1 \times .5) = .5$ vs. $(1 \times 1) + (0 \times .75) + (0 \times .5) = 1$. However, if the second candidate violates both C_2 and C_3, we get a different result:

(44)

	Input	C_1	C_2	C_3
	$cand_1$		*	*!
☞	$cand_2$	*		

Here we have $(0 \times 1) + (1 \times .75) + (1 \times .5) = 1.25$ vs. $(1 \times 1) + (0 \times .75) + (0 \times .5) = 1$.

This framework does not make specifically statistical predictions about the nature of the usual sort of phonological data, though it does, of course, make different typological predictions, which Pater (2009) explores. This kind of approach is developed statistically in several of the proposals reviewed below.

22.6 Phonotactics

Another domain in which statistical models have made an appearance in phonological theory is the treatment of corpus facts. An extremely influential early paper in this domain is Davis (1989, 1992).

Davis (1989) establishes on the basis of an examination of a computerized lexicon of English that certain otherwise expected sCVC sequences do not occur. Specifically, while identical voiceless stops can occur in CVC words (45), identical non-coronal voiceless stops cannot occur in sCVC words (46).

(45)

		C_2		
		p	t	k
	p	**pop**	pat	poke
C_1	t	tap	**tat**	tack
	k	cap	cat	**cake**

(46)

		C_2		
		p	t	k
	p	—	spit	speak
sC_1	t	step	**state**	stick
	k	skip	scat	—

Davis observes that this is most likely an OCP (Obligatory Contour Principle; Leben 1973) effect involving the marked place features labial and dorsal. What's striking is that the effect only seems to apply when the [s] is present on the left.

Berkley (1994a, b) establishes that statistically the restriction happens in both cases. That is, even in CVC contexts, the relevant word types, pVp and kVk, are underrepresented. Berkley breaks up the data based on the number of intervening segments; we'll only discuss the case where a single segment – a vowel – intervenes. In (47) we see the observed and expected distribution of voiceless stops across a single segment.

(47)

			C2			
		labial	coronal	dorsal		
	labial	**26**	256	42	observed
		64.9	204.5	54.6	expected
C1	coronal	204	**428**	160	observed
		158.7	**499.9**	133.5	expected
	dorsal	22	110	**10**	observed
		28.4	89.6	**23.9**	expected

As Berkley notes, the observed distribution is significantly different from chance: $X^2(4, N = 1258) = 81.971$, $p < .001$. This comparison is not quite the correct one. What we really want to do is compare the distribution for labial+labial and dorsal+dorsal to all other cells and ignore the differences within these categories.

(48)

lab+lab & dors+dors	All others	
36	1222	observed
279.55	978.44	expected

This more restrictive comparison is also significant: $X^2(1, N = 1258) = 272.818$, $p < .001$.

The upshot is that the categorical distribution Davis observes for sCVC is paralleled by a statistical skewing for CVC cases.

22.7 Experiments

The domain in which statistical phonology has been most aggressively applied is experimental psycholinguistics, performance measures that specifically address phonological questions. A very influential early paper in this area is Coleman & Pierrehumbert (1997). The paper proposes a statistical model for acceptability judgments built on dictionary data.

The model proposed was developed initially to model acceptability judgments (Coleman 1996). The model is a version of a probabilistic context-free grammar (Suppes 1970). The model is trained on monosyllables and disyllables in Mitton (1992). The key move in the construction of the model is that syllable stress and position are kept distinct throughout the parse tree.

First, the model distinguishes compounds from non-compounds.

(49) U → W
U → W W

Syllables are coded for whether they are stressed (Xs) or stressless (Xw), and for whether they are initial (Xxi), final (Xxf), or both (Xxif).

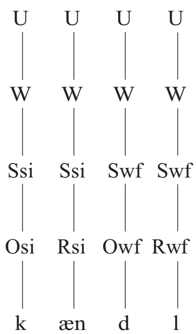
(50) W → Ssif
W → Swi Ssf
W → Ssi Swf
W → Ssi Ssf

Syllables branch into onsets and rhymes and stress and position information propagates downward.

(51) Ssif → Osif Rsif
Ssf → Osf Rsf
Ssi → Osi Rsi
Swi → Owi Rwi
Swf → Owf Rwf

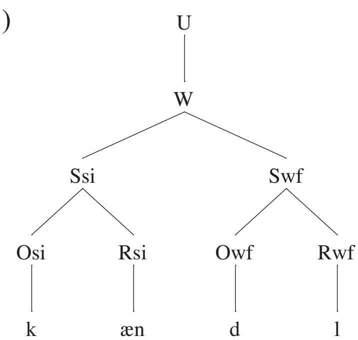
A word like *candle* would have paths like this:

(52)



Finally, a proper parse tree for a word like *candle* would look like this:

(53)



Coleman & Pierrehumbert trained their model on monosyllables and disyllables in Mitton (1992), and then the predicted probabilities for nonsense items were compared with the judgments from the acceptability task. Specifically, four different comparisons were made: (i) overall probability of the word; (ii) log probability of the word; (iii) probability of the worst part of the word; and (iv) probability of the best part of the word.

(54)

Scoring method	Significance of correlation
p(word)	$p < .01$
ln(p(word))	$p < .001$
p(worst part)	$p < .01$
p(best part)	n.s.

Significance testing was on the basis of a *t*-test on r^2. All methods reached significance except the best part of the word.

The point of these different comparisons was to get at the fact that the judgment task seemed to be sensitive to global probabilities rather than constraint violations per se. Consider a nonsense item like *mrup* [mrup]. We would expect this to be less acceptable because of the illegal Onset [mr]. Interestingly, however, a form like *mrupation* [mrupeʃən] with the same illegal Onset is judged as *more* acceptable. This comparison suggests that acceptability judgments are a function of the probabilities of all the elements in the item, not just of elements that violate some categorical phonological restriction (e.g. on Onset structure). Incidentally, this comparison also argues against a transitional probability or *N*-gram-type approach as, under these approaches, a form like *mrupation* could not have a higher probability than *mrup*, which forms a proper prefix of the former. (See Hammond 2003 for a similar approach.)

Pierrehumbert (1994) presents a different statistically based approach to well-formedness. Specifically, Pierrehumbert asks whether the well-formedness of medial clusters is a function of the statistical distribution of clusters at word edge. The logic is that we expect medial clusters to be the result of a legal syllable-final cluster being juxtaposed with a legal syllable-initial cluster. (See Hammond 1999b for a general presentation of these clusters and a non-statistical treatment.) Using an electronic dictionary, Pierrehumbert shows that there is a good fit between the predicted frequency of medial clusters and their actual frequency based on peripheral cluster frequencies. She confirms some of the patterns observed with a well-formedness experiment. This effect supports again the idea that well-formedness can be viewed as a statistical notion.

In an experimental study, Coetzee (2008) also investigates the sCVC restrictions discovered by Davis, arguing that "grammar in addition to frequency statistics influences processing" (p. 228). His general proposal is very similar to that of Golston: that we can order OT candidates in a more detailed fashion than winner vs. loser and that the number and rank of constraint violations corresponds to well-formedness. Let's look at this in more detail.

The key question with respect to the sCVC facts is whether the apparent restriction against spVp and skVk is accidental. Coetzee proposes a grammatical organization and learning algorithm where they must be part of the grammar. He argues that processing facts support this conclusion.

The basic organization of grammar he proposes is that there is a set of derived constraints against spVp, skVk, and stVt. These constraints are derived via local conjunction from more basic constraints against s+stop sequences and OCP violations.

> English tolerates violation of each of the constraints . . . individually – *toot* (*tVt), *cake* (*kVk), *pope* (*pVp), *sky* (*[s + stop]$_\sigma$), *whisk* (*[s + stop]$_\sigma$), and so forth. What English does not tolerate, however, is the violation of some combinations of these constraints – English does not tolerate violation of *[s + stop]$_\sigma$ together with either *kVk or *pVp.
>
> *(Coetzee 2008: 232)*

As far as ranking these derived, locally conjoined constraints, since violations of *stVt occur, it follows that that constraint is ranked below relevant faithfulness constraints and the other two above.

(55) {*spVp, *skVk} ≫ Faithfulness ≫ *stVt

Such a ranking makes exactly the same predictions as the same ranking without a constraint against *stVt, but Coetzee argues that the latter is necessary. In addition, Coetzee argues that rankings are learned partially on the basis of frequency distributions (Pater 2005). He cites the following distributional patterns from the CELEX dictionary:

(56)

Syllable type	Count per million
$_\sigma$[. . . tVt . . .]$_\sigma$	5348
$_\sigma$[. . . kVk . . .]$_\sigma$	695
$_\sigma$[. . . pVp . . .]$_\sigma$	235

If frequency affects rankings, then these facts suggest that the full ranking of these constraints is:

(57) *spVp ≫ *skVk ≫ Faithfulness ≫ *stVt

Coetzee conducts several experiments to test this analysis. (See Hammond 2012 for a discussion of the general relationship between corpus and experimental data.) First, there is a phoneme identification experiment. He constructed three acoustic continua for place of articulation among stop consonants. We know that overall transitional probabilities and neighborhood density can have effects in this domain, so Coetzee controls for these in his design and subsequent statistics. His grammatical model predicts the biases diagrammed below.

(58)

Condition	Continuum	Expected bias	Constraint
K ~ P	[spαp] ~ [spαk]	against [p]	*spVp
	[skαp] ~ [skαk]	against [k]	*skVk
T ~ K	[skεk] ~ [skεt]	against [k]	*skVk
	[stεk] ~ [stεt]	toward [k]	*stVt
T ~ P	[spʌp] ~ [spʌt]	against [p]	*spVp
	[stʌp] ~ [stʌt]	toward [k]	*stVt

For example, we expect a bias toward [k] in the [stεk] ~ [stεt] continuum if there is a constraint against *stVt. Likewise, we expect a bias against [k] in the [skαp] ~ [skαk] continuum if there is a constraint against *skVk. Notice that these predictions allow us to test both novel aspects of the analysis. First, we can test whether there is indeed a constraint against *stVt. Second, we can test the proposed ranking of *spVp above *skVk.

A total of thirty-seven subjects took part. In each condition only those subjects who could identify the endpoint stimuli correctly at least 75% of the time were included in the analysis. This resulted in fifteen subjects in the K ~ P condition, twenty-six in the K ~ T condition, and twenty-six in the P ~ T condition. Coetzee shows that all of the expected effects occur.

The next experiment was a word-likeness experiment that directly compared the different OCP violations: [stVt] ~ [skVk], [stVt] ~ [spVp], and [skVk] ~ [spVp]. Again, transitional probabilities and neighborhood density can affect responses, so again these were controlled for. In all cases, word-likeness judgments supported the ranking given in (57).

Finally, Coetzee also conducted a lexical decision experiment. Here subjects were presented nonsense items in frame sentences. The relevant question is how quickly subjects would reject items that violated one of the sCVC constraints, with the prediction that non-word items that violated a higher-ranked OCP constraint would be rejected more quickly than one that violated a lower-ranked OCP constraint. Using ANOVA, this effect was significant by subjects, but not by items.

The upshot is that behavioral measures largely converge with the predictions of a grammar where constraint ranking is a partial function of frequency, e.g. *spVp ≫ *skVk. In addition, a constraint motivated by general phonological terms and not language-specific data, *stVt, is also needed.

22.8 Lexical frequency

Coetzee & Kawahara (2013) develop a different statistical approach to the role of lexical frequency in variationist data that they test against coronal deletion in English and geminate devoicing in English borrowings in Japanese.

It's long been known that lexical frequency affects the likelihood of some phonological processes. For example, Fidelholtz (1975) cites the contrast between *astronomy* [əstránəmi] and *gastronomy* [gæ̀stránəmi] as evidence that pretonic vowel reduction in English is sensitive to lexical frequency; vowel reduction is much more likely in the more frequent *astronomy* than in the less frequent *gastronomy*. Hooper (1976) treats a similar effect with respect to vowel syncope. In English, a post-tonic vowel is more likely to syncopate in a high-frequency word, e.g. *memory* [mɛ́mri], than in a low-frequency word, e.g. *mammory* [mǽməri].

Hammond (1999a) treats this effect in OT terms. His data come from the English Rhythm Rule (Liberman & Prince 1977; Hayes 1984; Hammond 1988; etc.). Specifically, he shows that rhythmic stress shift in phrases is more likely when the first word is more frequent, e.g. *ábstràct plán* vs. *àbstrúse plán*. (Hicks et al. 2000 go on to show similar effects based on phrasal frequency or likelihood; Hammond 2004 treats word-internal morpheme-level frequency effects.) This effect is treated by positing lexically specific constraints that separate frequent and infrequent words. For example:

(59) Faith(àbstrúse) ≫ Rhythm ≫ Faith(àbstráct)

A similar approach is taken by Pater (2000). There are several disadvantages of this general approach. First it is imprecise; how do we connect specific frequencies to the ranking? Second, the approach relies on multiple faithfulness constraints, one for each lexical item. This flies in the face of the universality of the OT constraint set.

Coetzee & Kawahara (2013) develop an approach that addresses these concerns. First they make use of the Harmonic Grammar framework that we've already discussed with its system of finite weights; however, they augment that framework with evaluation noise as in Stochastic OT. Recall the general logic for Harmonic Grammar:

(60) $$H(cand) = \sum_{i=1}^{n} w_i C_i(cand)$$

where w_i is the weight of constraint C_i, and C_i(*cand*) is the number of violations of candidate *cand* in terms of C_i expressed as a negative integer.

In *Noisy Harmonic Grammar*, a noise value is added to the weight of each constraint each time that the grammar is used. Coetzee & Kawahara cite the following schematic example to show how it works. What we have is a case where a consonant cluster may or may not be simplified depending on the variable ranking that noisy evaluation allows.

(61)

	w nz 5 −0.7	w nz 1.5 −0.4	w nz 1 0.2	
/lʊst/	Dep (4.3)	*Complex (1.1)	Max (1.2)	H
☞ lʊst		−1		−1.1
lʊs			−1	−1.2
lʊs.ti	−1			−4.3

	w nz 5 −0.7	w nz 1.5 0.1	w nz 1 −0.1	
/lʊst/	Dep (4.3)	*Complex (1.1)	Max (1.2)	H
lʊst		−1		−1.6
☞ lʊs			−1	−0.9
lʊs.ti	−1			−4.3

On this view, we alter the definition of harmony to include a term *nz* for evaluation noise:

(62) $$H(cand) = \sum_{i=1}^{n}(w_i + nz_i)C_i(cand)$$

where w_i is the weight of constraint C_i, nz_i the noise associated with constraint C_i at this evaluation occasion, and $C_i(cand)$ is the number of violations of candidate *cand* in terms of C_i expressed as a negative integer.

Coetzee & Kawahara propose to augment this framework with an additional term to accommodate lexical frequency. Specifically, at evaluation faithfulness constraints include an additional scaling factor *sf* depending on the lexical frequency of the item under evaluation, such that faithfulness constraints are effectively weighted more for infrequent words and weighted less for frequent words.

Imagine we have a fairly frequent item /lʊst/ with *sf* = −1.

(63)

	w nz sf 5 −0.7 −1	w nz 1.5 0.1	w nz sf 1 0.2 −1	
/lʊst/	Dep (4.7)	*Comp (1.6)	Max (0.2)	H
lʊst		−1		−1.6
☞ lʊs			−1	−0.2
lʊs.ti	−1			−4.7

We can compare this with a less frequent item /nʊst/ with *sf* = +1.

(64)

	w nz sf 5 −0.7 1	w nz 1.5 0.1	w nz sf 1 0.2 1	
/nʊst/	Dep (4.7)	*Comp (1.6)	Max (0.2)	H
☞ nʊst		−1		−1.6
nʊs			−1	−2.2
nʊs.ti	−1			−6.7

We then alter the definition of harmony to include the scaling factor for faithfulness constraints as follows:

(65) $$H(cand) = \sum_{i=1}^{n}(w_i + nz_i)M_i(cand) + \sum_{j=1}^{m}(w_j + nz_j + sf)F_j(cand)$$

where M_i is the *i*-th markedness constraint, w_i is the weight associated with M_i, nz_i the noise associated with M_i at this evaluation occasion, and $M_i(cand)$ the number of violations of candidate *cand* in terms of M_i expressed as a negative integer; and where F_j is the *j*-th faithfulness constraint, w_j the weight associated with F_j, nz_j the noise associated with F_j at this evaluation occasion, and $F_j(cand)$ the number of violations of candidate *cand* in terms of F_j expressed as a negative integer; and where *sf* is the scaling factor associated with the specific word being evaluated. Notice how, on this approach, there is no multiplication of faithfulness constraints; we simply allow those constraints access to the scaling factor for each item.

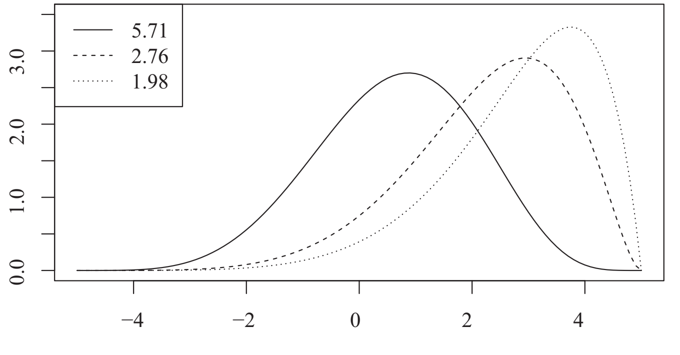
Where do the scaling factors come from? Coetzee & Kawahara propose that they come from a form of the *beta* distribution:

(66) $$f(x,\alpha,\beta,\rho) = \rho \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1}(1-x)dx}$$

The basic logic of the distribution in its usual form is that it looks very much like the normal distribution, but is bounded in the range [0,1], unlike the normal. The intent of this version of the distribution is that the range be centered on zero and extendable to larger intervals based on values for ρ. The other parameters of the distribution, α and β, control the skew of the distribution here as with the usual form of the distribution.

Coetzee & Kawahara take α as the median log frequency for the sampled lexical items. The value of β is determined based on the log frequency of the specific lexical items. Finally, ρ is set based on the global fit of the model. The net effect is that the scaling factor for infrequent words is negative and for frequent words is positive. Here's an example of plotted beta distributions based on a reference level (α) of 5.1, β values of {5.71, 2.76, 1.98}, and ρ = 5.

(67) *Beta* distributions for words with ρ = 5



The scaling factor is actually set to the mode of the distribution, so most of the distribution is irrelevant, and most of the math can be consequently simplified. The mode of a beta distribution is:

(68) $$\frac{\alpha - 1}{\alpha + \beta - 2}$$

We can then calculate the mode of the relativized distribution like this:

(69) $$\left(\frac{\alpha - 1}{\alpha + \beta - 2} \times 2\rho\right) - \rho$$

Coetzee & Kawahara test this theory against the facts of coronal deletion in English and geminate devoicing in English borrowings in Japanese. For the English data, they collect relevant examples from the Buckeye corpus (Pitt et al. 2007). To simplify the mapping, they separate items into ten bins by log frequency. The bins have the log frequency values shown below.

(70)

Frequency bins	Baseline	ρ = 3	4	5	6	7
2	0.0	1.82	2.43	3.04	3.65	4.25
2.6	0.0	1.32	1.75	2.19	2.63	3.07
3	0.0	1.03	1.38	1.72	2.07	2.41
3.5	0.0	0.73	0.97	1.21	1.45	1.7
4	0.0	0.46	0.62	0.77	0.93	1.08
4.4	0.0	0.28	0.37	0.47	0.56	0.65
5.8	0.0	−0.24	−0.31	−0.39	−0.47	−0.55

(Values in the ρ columns: Scaling factor)

They fit the model with different values for ρ, eventually setting on $\rho = 5$.

Coetzee & Kawahara show a good fit between the model developed and the English and Japanese data. The model is clearly a step forward in the treatment of lexical frequency as it offers a statistically precise treatment. That said, it's not clear that the beta distribution per se is required. In addition, while they can draw on the general learning mechanisms for Harmonic Grammar, they must still stipulate that frequency effectively alters the ranking of faithfulness constraints.

22.9 Maximum entropy

Hayes & Wilson (2008) develop a statistical model of phonotactics and an algorithm for learning phonotactic grammars. They test the theory they develop on English onsets, Shona vowel harmony, stress systems, and the full phonology of Wargamay. (See Goldwater & Johnson 2003 for the general notion of a maximum entropy OT grammar.)

The model of grammar Hayes & Wilson develop is similar to the finitely weighted versions of OT we've looked at already. The *score* of a phonological representation x, denoted h(x), is given as:

(71) $$\mathrm{h}(x) = \sum_{i=1}^{N} w_i C_i(x)$$

where, for any C_i in (C_1, C_2, . . . , C_N): $w_i \geq 0$.

These scores are then converted to *maxent values*. Given a phonological representation x and its score h(x) under a grammar, the maxent value of x, denoted P*(x), is:

(72) $P^*(x) = \exp(-h(x))$

Hayes & Wilson cite the following schematic example. Here we have three different CV shapes and two constraints, one against word-initial vowels and another against word-final consonants. Given the weights shown, the score for each representation is given in the fourth column, and the maxent value in the fifth column.

(73)

x	*#V $w = 3$	*C# $w = 2$	Score H(x)	Maxent value P*(x)
CV	3×0	2×0	$0+0=0$	$\exp(-0) = 1$
CVC	3×0	2×1	$0+2=2$	$\exp(-2) = .14$
V	3×1	2×0	$3+0=3$	$\exp(-3) = .05$

Hayes & Wilson then convert maxent values to probabilities. To do this, they must relativize maxent values to the set of possible values. If we take **Ω** to be the (infinite) set of all universally possible phonological surface forms, we can define the probability of a specific maxent value as follows.

(74) $$P(x) = \frac{P^*(x)}{\sum_{y \in \Omega} P^*(y)}$$

Given a maxent grammar and a set D of observed data, the probability of D under the grammar is:

(75) $$P(D) = \prod_{x \in D} P(x)$$

After converting this to a log, Hayes & Wilson take the learning problem to be to maximize log(P(D)), that is, find the grammar that makes this value the biggest it can be.

This is only half the problem, however. The logic above will allow us to use known techniques from machine learning to find the best weighting for some set of constraints, but it does not find the best constraints. Hayes & Wilson suppose that the maxent phonology of a language is not formed just by weighting a universal set of constraints, but by choosing a subset of those constraints and weighting them.

They propose that constraints all exhibit a common format: a constraint is a finite sequence of partial feature matrices:

(76) $$\begin{bmatrix} \alpha F \\ \beta G \\ . \\ . \\ . \end{bmatrix} \begin{bmatrix} \gamma H \\ \delta I \\ . \\ . \\ . \end{bmatrix} \cdots \begin{bmatrix} \varepsilon J \\ \zeta K \\ . \\ . \\ . \end{bmatrix}$$

One of the matrices of a constraint can contain the complementation operator; thus [ˆαF, βG, . . .] means any segment not in the natural class [αF, βG, . . .]. If the number of features and feature values is finite, and the number of matrices that can figure in any one constraint is bounded at some finite point, then the number of possible constraints is finite.

Finding the right constraints is done by traversing the finite space of possible constraints and looking for the subset that offers the best fit to the data (as described above). Two additional principles govern the selection of constraints: accuracy and generality.

(77) Accuracy
Select constraints that are violated as little as possible by the actual data ($O[C_i]$) as compared with the possible data ($E[C_i]$).

(78) Generality
Prefer constraints with fewer matrices and where matrices define more general classes.

The general algorithm is as follows, basically using accuracy and generality to guide the search process:

(79) Phonotactic learning algorithm
Input: a set Σ of segments classified by a set $\mathcal{F}$ of features, a set $\mathcal{D}$ of surface forms drawn from Σ^*, an ascending set $\mathcal{A}$ of accuracy levels, and a maximum constraint size $\mathcal{N}$.

1: begin with an empty grammar $\mathcal{G}$
2: **for** each accuracy level a in $\mathcal{A}$ **do**
3:     **repeat**
4:         select the most general constraint with accuracy less than a (if one exists) and add it to $\mathcal{G}$
5:         train the weights of the constraints in $\mathcal{G}$
6:     **while** a constraint is selected in step 4
7: **end for**

Hayes & Wilson test their system against a range of data, both experimental and corpus based, showing an excellent fit in all cases. This is a huge step forward, but it should not be all that surprising. The system finds the best constraints and the best fit because it basically considers the best fit for every possible constraint. It would be more surprising if this system did *not* find an excellent fit overall. (The Hayes & Wilson approach is an instance of a more general approach to constraint weighting in OT that uses maxent. See Coetzee & Pater 2011 for discussion of how that more general system fares with respect to other weighting schemes.)

That said, the maxent approach has set a new benchmark in statistical phonological approaches. We can achieve excellent results with unconstrained weights and a full traversal of the constraint search space. We must now ask whether this kind of searching + weighting scheme corresponds to what people actually do. Does the learner consider every possible option as described here? Does the learner exhibit incremental grammars that reflect the steps we would expect from (79)?

22.10 Conclusion

We've seen that phonological theory from the very outset of generative grammar has been challenged by quantitative non-traditional data. These include sociolinguistic data, corpus data, including frequency effects, and experimental data from various sources, but most especially well-formedness data.

These data have necessitated revision to the basic architecture of the theory, and these revisions have recurred as the architecture of the theory has moved from rules to constraints. Thus, while we had variable rules under classical rule-based phonology, we have analogous machinery under modern constraint-based approaches.

Over the years, the approaches have varied in terms of how quantitative effects have been incorporated. Some early approaches have been understandably thin in terms of statistical sophistication, but the bar has been steadily raised over time so that now work in this area is quite sophisticated mathematically and computationally.

What can we expect in terms of the future? It's apparent that the range of new data will increase, both qualitatively and quantitatively. Thus we expect different kinds of corpora and different kinds of experimental techniques will be treated in statistical phonological terms. We also expect larger and larger datasets and corpora to be treated. It's unclear whether changes in scale will have theoretical consequences, but they will certainly change the quality of the statistical models that can be developed.

Finally, we expect that the statistical and mathematical theories invoked will become richer. This will occur both in terms of increased statistical expertise on the part of phonologists, but also in terms of the fields we can borrow techniques from, e.g. machine learning, statistical natural language processing, etc.

While we have seen that we can extend phonology to accommodate non-traditional data, the real question is whether the theories we develop for corpus data, experimental data, etc. will in turn tell us more about traditional judgment data.

22.11 Further reading

Labov (1969) is the classic presentation of the variable rule formalism. Boersma (1997) is the original paper on Stochastic Optimality Theory. Davis (1989) is the first discussion of the CVC/sCVC restrictions in English, and Coetzee (2008) presents experiments that test the CVC/sCVC restrictions and extend the pattern. Finally, Hayes & Wilson (2008) is the first presentation of maxent phonology.

22.12 Related topics

- Rule-based phonology
- Optimality Theory
- Connectionist phonology
- Exemplar theory
- Laboratory phonology

Note

1 Thanks to Shannon Grippando, an anonymous reviewer, and the editors. All errors are my own.

References

Anttila, Arto (1995). Deriving variation from grammar: A study of Finnish genitives. ROA #63.

Berkley, Deborah Milam (1994a). The OCP and gradient data. *Studies in the Linguistic Sciences* 24:59–72.

Berkley, Deborah Milam (1994b). Variability and Obligatory Contour Principle effects. In Katherine Beals et al. (eds.), *CLS 30: Proceedings of the 30th Annual Meeting of the Chicago Linguistic Society, Vol. 2: The Parasession on Variation and Linguistic Theory*. Chicago: Chicago Linguistic Society, 1–12.

Boersma, Paul (1997). How we learn variation, optionality, and probability. ROA #221.

Boersma, Paul & Bruce Hayes (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Cedergren, Henrietta & David Sankoff (1974). Variables rules: Performance as a statistical reflection of competence. *Language* 50:333–355.

Chomsky, Noam (1957). *Syntactic Structures*. The Hague/Paris: Mouton.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Coetzee, Andries W. (2008). Grammaticality and ungrammaticality in phonology. *Language* 84:218–257.

Coetzee, Andries W. & Shigeto Kawahara (2013). Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 31:47–89.

Coetzee, Andries & Joe Pater (2011). The place of variation in phonological theory. In John Goldsmith, Jason Riggle & Alan Yu (eds.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 401–434. ROA #946.

Coleman, John (1996). The psychological reality of language-specific constraints. Paper presented at the Fourth Phonology Meeting, University of Manchester.

Coleman, John & Janet Pierrehumbert (1997). Stochastic phonological grammars and acceptability. In *Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology*. Somerset: Association for Computational Linguistics, 49–56.

Davis, Stuart (1989). Cross-vowel phonotactic constraints. *Computational Linguistics* 15:109–111.

Davis, Stuart (1992). Investigating English phonotactic constraints using a computerized lexicon. In Ursula Klenk (ed.), *Computatio Linguae: Aufsätze zur Algorithmischen und Quantitativen Analyse der Sprache*. Stuttgart: Franz Steiner, 1–15.

Fidelholtz, James (1975). Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11:200–213.

Frisch, S., N. R. Large & D. B. Pisoni (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Langauge* 42:481–496.

Goldsmith, John (1993). Harmonic phonology. In John Goldsmith (ed.), *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago: University of Chicago Press, 21–60.

Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson & Osten Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University, 111–120.

Golston, Chris (1998). Constraint-based metrics. *Natural Language and Linguistic Theory* 16:719–770.

Greenberg, Joseph H. (ed.) (1978). *Universals of Human Language: Phonology*, vol. 2. Stanford: Stanford University Press.

Greenberg, Joseph H. & J. J. Jenkins (1964). Studies in the psychological correlates of the sound system of American English. *Word* 20:157–177.

Guy, Gregory R. (1991). Explanation in a variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3:1–22.

Hammond, Michael (1988). *Constraining Metrical Theory: A Modular Theory of Rhythm and Destressing*. New York: Garland. 1984 UCLA doctoral dissertation.
Hammond, Michael (1994). An OT account of variability in Walmatjari stress. ROA #20.
Hammond, Michael (1999a). Lexical frequency and rhythm. In Michael Darnell et al. (eds.), *Functionalism and Formalism in Linguistics*. Amsterdam: John Benjamins, 329–358.
Hammond, Michael (1999b). *The Phonology of English*. Oxford: Oxford University Press.
Hammond, Michael (2003). Phonotactics and probabilistic ranking. In Heidi Harley, Andrew Carnie & Mary Willie (eds.), *Formal Approaches to Function in Grammar: In Honor of Eloise Jelinek*. Amsterdam: Benjamins, 319–332.
Hammond, Michael (2004). Frequency, cyclicity, and optimality. *Studies in Phonetics, Phonology, and Morphology* 10:349–364.
Hammond, Michael (2012). Corpus data vs. experiments in English phonotactics. In James Myers (ed.), *In Search of Grammar: Experimental and Corpus-based Studies*. Taipei: Academia Sinica, 215–240.
Hayes, Bruce (1984). The phonology of rhythm in English. *Linguistic Inquiry* 15:33–74.
Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379–440.
Hicks, Cathy, Robert Kennedy & Michael Hammond (2000). Lexical frequency and the rhythm rule in English. Annual meeting of the LSA.
Hooper, Joan (1976). Word frequency in lexical diffusion and the source of morphophonological change. In William M. Christie (ed.), *Current Progress in Historical Linguistics*. Amsterdam: North-Holland, 96–105.
Itô, Junko (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7:217–260.
Keller, Frank (2006). Linear Optimality Theory as a model of gradience in grammar. In Gisbert Fanselow, Caroline Féry, Ralph Vogel & Matthias Schlesewsky (eds.), *Gradience in Grammar: Generative Perspectives*. Oxford: Oxford University Press, 270–287.
Kiparsky, Paul (1982). Lexical morphology and phonology. In In-Seok Yang (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin, 3–91.
Kisseberth, Charles (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1:291–306.
Labov, William (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.
Leben, William (1973). *Suprasegmental Phonology*. Ph.D. thesis, MIT.
Legendre, Geraldine, Yoshiro Miyata & Paul Smolensky (1990a). Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: An application. *Technical report* #90–4. Institute for Cognitive Science, University of Colorado at Boulder.
Legendre, Geraldine, Yoshiro Miyata & Paul Smolensky (1990b). Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. *Technical report* #90–5. Institute for Cognitive Science, University of Colorado at Boulder.
Liberman, Mark & Alan Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* 8:249–336.
Maye, Jessica (2000). *The Acquisition of Speech Sound Categories on the Basis of Distributional Information*. Ph.D. thesis, University of Arizona.
McCarthy, John & Alan Prince (1993). Prosodic morphology. U. Mass.
Mitton, R. (1992). A computer-usable dictionary file based on the oxford advanced learner's dictionary of current English. http://ota.ox.ac.uk.
Ohala, John & Manjari Ohala (1986). Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In John Ohala & Jeri Jaeger (eds.), *Experimental Phonology*. Orlando: Academic Press, 239–252.
Pater, Joe (2000). Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17:237–274.

Pater, Joe (2005). Learning a stratified grammar. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *Proceedings of the 29th Boston University Conference on Language Development*. Somerville: Cascadilla, 482–492.
Pater, Joe (2009). Weighted constraints in generative linguistics. *Cognitive Science* 33:999–1035.
Pierrehumbert, Janet B. (1994). Syllable structure and word structure. In Patricia Keating (ed.), *Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press, 168–190.
Pitt, Mark A., Laura Dilley, Keith Johnson, Scott Kiesling, William Raymond, Elizabeth Hume & Eric Fosler-Lussier (2007). Buckeye corpus of conversational speech (2nd release). www.buckeyecorpus.osu.edu.
Potts, Christopher, Joe Pater, Karen Jesney & Rajesh Bhatt (2010). Harmonic Grammar with linear programming: From linear systems to linguistic typology. *Phonology* 27:77–117.
Prince, Alan & Paul Smolensky (1993). Optimality Theory. University Mass and University of Colorado.
Sankoff, David & William Labov (1979). On the uses of variable rules. *Language in Society* 8:189–222.
Suppes, Patrick (1970). Probabilistic grammars for natural languages. *Synthèse* 22:95–116.
Turton, Danielle (2012). The darkening of English /l/: A Stochastic Stratal OT analysis. http://ling.auf.net/lingbuzz/001524.
Zamuner, Tania (2001). *Input-Based Phonological Acquisition*. Ph.D. thesis, University of Arizona.
Zamuner, Tania, LouAnn Gerken & Michael Hammond (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language* 31:515–536.

23

Phonology and evolution

Bart de Boer

23.1 The basics of evolution

Evolution, reduced to its bare essentials, consists of three rather abstract processes. These are: (1) transmission of information (which implies storage of information as well), (2) variation on this information and (3) selection of certain variants over others. In practical instances of evolution, these different processes take different forms. In the best-known instance of evolution – biological evolution – information is stored in the chromosomes of a cell. The chromosomes consist of DNA. Variation is caused by mutation of this DNA and through recombination and crossover of the chromosomes. Selection happens because the DNA determines how an organism grows and functions. As certain variations will cause organisms to function better, those variations will have more surviving offspring on average than less good variations. By definition, the number of offspring in the next generation is called the fitness of that organism. Finally, transmission happens because DNA is copied relatively faithfully and transmitted to the offspring of an organism. Given these definitions, it follows (mathematically) that variants that have higher fitness will tend to occur more frequently in the next generation than in the present generation. Therefore evolution causes a population of organisms to adapt to the environment in which it lives.

Because transmission, variation and selection are entirely general processes, evolution can occur in different ways than in the biological example. Another example is cultural evolution. Here learning of observed behavior causes transmission, variation consists of conscious or subconscious modifications when reproducing the behavior, and selection depends on which kinds of behavior are most readily learned and reproduced by others. The constraints of transmission, variation and selection are very different. For instance, information can be transmitted to anyone, not just one's direct offspring, and it can be received from many sources, not just from one or two parents.

Variation is not just a random process, but can involve conscious modification. Selection can similarly happen through conscious choice, rather than through the relatively random processes of biological evolution. Nevertheless, cultural evolution can cause the emergence, spread and complexification of learned behavior similar to the way in which biological evolution can cause the emergence, spread and complexification of physical adaptations.

Language is a culturally transmitted behavior and therefore prone to cultural evolution. In addition, physical adaptations to language are prone to biological evolution, because language is an important part of the human environment and a person's reproductive success depends in part on how well they can communicate. Phonology is the aspect of language for which both evolutionary processes are perhaps easier to observe than for any other aspects of language. Because phonology is the aspect of language that is closest to the physical signal, it is possible to identify anatomical adaptations that deal with it: the ones involved in producing and perceiving signals. These adaptations can then also be studied in other animals and in fossils. As for cultural evolution, the study of historical language change gives a lot of precise information on how speech changes over the generations. Moreover, phonology is an aspect of language that is acquired early and that is difficult to modify consciously later in life. Therefore its transmission and variation are somewhat closer to those in biological evolution (as will be elaborated in section 23.4 on language change below), and therefore it is easier to study and model than cultural evolution of aspects of behavior that are learned in a more conscious way.

Because language is a culturally evolving system and human language users evolve biologically, language is the result of co-evolution between these two evolving systems. Co-evolution happens when two evolving systems influence each other's fitness. Co-evolution is at the center of recent debates on the evolution of language. It is therefore useful to recapitulate a classic (but highly idealized) example from biology: that of the cheetah and the gazelle. Cheetahs need to eat gazelles in order to reproduce, and therefore cheetahs that are more successful in running down gazelles produce more offspring. Thus there is a selective pressure for cheetahs to be faster than gazelles. Note however that the cheetah's fitness is in part determined by the behavior of the gazelles. This behavior is also under selective pressure: gazelles that live longer produce more offspring, and longevity is determined by how well they can outrun cheetahs. Therefore there is a selective pressure for gazelles to be faster than cheetahs. Thus, the fitness of gazelles is also partly determined by the behavior of cheetahs. Interaction between two evolving systems leads to interesting effects. In the case of the cheetah and the gazelle it may have caused them to become ever faster. As will be discussed in section 23.5 on the interaction between culture and biology, in language the effects may be even more complex.

Part of this complexity is due to culture and biology evolving at different time scales: language change (which can be seen as an instance of cultural evolution) occurs at a time scale of hundreds to thousands of years, whereas biological evolution happens at the time scale of tens of thousands to hundreds of thousands of years. Because language evolves more rapidly than biology, it provides a "moving target". Biological adaptations can therefore not evolve for all aspects of language. This problem is at the heart of the debate on the role of culture–biology co-evolution that will be discussed in more detail below.

Cultural evolution is much faster than biological evolution for several reasons. First of all, the period of time between transmissions of information is not limited to the time between generations. In principle, every time someone makes an utterance, cultural information is transmitted. Secondly, one utterance can influence multiple listeners and this also means that cultural information can spread more rapidly. Thirdly, variation of information need not be as random as variation in biological evolution. It is possible that speakers consciously or subconsciously control the variation of their utterances, preferring variations that have a higher probability of being selected by the listeners. Finally, selection of linguistic variants is also less prone to random factors than in biology. Listeners can make informed decisions about which variants they prefer.

Incidentally, these differences between the precise way in which transmission, variation and selection work also help to understand the differences between different kinds of cultural evolution. For instance, it can help to understand why (cultural) evolution of technology can proceed much faster than that of language. First of all, transmission can be even more widespread for technology, because it can also be transmitted through the artifacts it creates. Transmission of technology is therefore not limited to the number of people one can talk to, nor is it limited by critical period effects that occur in language learning. Variations can be made consciously, especially if one understands the principles behind the technology, and finally, selection can also be made very effectively on the basis of which technologies work better than others. However if transmission is limited (because of isolation), and variation is too risky (because subsistence is only marginal), and selection tends to favor existing variants over new variants because of distrust of new (and therefore risky) technology, technological evolution can also be very slow. Because transmission, variation and selection work differently for different aspects of language, understanding their effects may help to understand why different aspects of language change over time in different ways.

For instance, aspects of language that are learned very early in life, and that are difficult to change later in life (such as phonology or inflexional morphology), tend to behave more like biological systems, with relatively limited transmission, variation mostly due to random and functional factors and selection mostly due to functional factors and frequency of occurrence. On the other hand, aspects of language that are used more productively and creatively (such as proverbs and other multi-word expressions) will evolve much more like technology, with potentially very wide transmission and conscious, non-random variation, while selection is based on what impresses listeners most.

23.2 The effect of evolution

The effect of evolution is almost tautologically to optimize fitness in the population. Evolution can thus be considered a process that optimizes a population for some purpose. In biology this translates to organisms becoming better adapted to their environment. In language it means that language becomes more easily learnable and more usable (in the broadest sense of the word). However, things are complicated by the fact that usability means different things to speakers than to listeners. Things are also complicated by the fact that learnability and usability of any aspect of language can depend on other aspects. For instance, a word like [ia] meaning "yes" in a language would be perfectly usable (easy to pronounce, and relatively robust to noise for listeners) unless there would be a very similar word, for instance [ea], meaning "no". Therefore optimization in language is not a straightforward process, and there are many satisfactory solutions (witnessed by the many different languages and dialects that exist or have existed). However, extremely suboptimal language should not exist, and although some subsystems of language (such as the vowel system) may be suboptimal, evolutionary theory would predict that suboptimal systems will not be stable over a long period.

It should be stressed that although evolution optimizes success in reproduction, it is not teleological, i.e. it does not work towards a goal. It is a purely local process in that it is only concerned with individual behavior (even though its effects are often described at the population level) and only concerned with what happens *now* – it cannot make use of information from the past and has no notion of what will happen in the future. The end effect of evolution may appear to be goal-oriented (i.e. bigger brains or vocal tracts that are better for speech) but those are the result of selection and variation (and transmission of course) over a longer period of time, in a relatively stable environment (i.e. the properties that cause higher reproductive success remain constant over time). When co-evolution takes place, or in a rapidly changing environment, evolution may appear much less goal-oriented.

The effectiveness with which evolution converges to optimal solutions also depends on the tradeoff between variation and selection. Selection causes certain solutions (be they biological adaptations or linguistic variants) to become more prevalent. Nevertheless, if this process continues for a long time and no new variants are introduced, the population will become uniform and evolution stops. Variation must therefore be introduced continuously for evolution to keep working. Variation is generally introduced by random (or consciously introduced) modification of existing information. However, this means that occasionally an existing good variant may be modified. If there is too much variation, existing good variants may disappear from the population. For evolution to result in continued improvement, a fine tradeoff between selection and variation is therefore necessary.

The tradeoff between variation and selection helps to understand when cultural evolution is likely to occur and when it is not. For instance, if an organism can learn behavior, but only in a very limited range, there is not enough variation for evolution to work on. This may perhaps explain the lack of cultural evolution in chimpanzee vocalizations: although they apparently can adapt their calls a little bit, the extent is very limited. On the other hand, if the learning process is very inaccurate, evolution also does not take place, as the behavior in the next generation is too different from that in the previous generation. This may help to explain why chimpanzee material culture does not undergo evolution: basically each individual reinvents culturally transmitted behaviors (such as nut cracking), which causes too much variation for evolution to work: transmission is not sufficiently accurate to copy slight differences in behavior that make the difference between adequate and excellent nut cracking. In human material culture, copying is much more accurate so that variation does not swamp slight selective pressures. Even slight improvements get transmitted.

Because evolution only works when variation is limited, it can only work by modifying things that are already there. Radically new solutions rarely, if ever, appear. Innovation tends to happen when existing structures are modified to fulfill a different function. This is called exaptation. Exaptation of a structure for a new behavior is in general only possible when (some of) the previously existing functions of that structure become less critical.

23.3 Evolution of anatomical adaptations to speech

The study of the evolution of speech investigates which traits of human anatomy and cognition have undergone selective pressure related to language. These are the traits of human anatomy that have changed over time to improve our ability to produce speech. Similarly, in the study of the evolution of (possibly signed) language, one investigates those traits of anatomy and cognition that have undergone selection related to language. Note that speaking in terms of “language or speech-specific” versus “domain general” should be avoided. This is because (as explained above) evolution needs to work with what is already there, and therefore any adaptation to relatively recent behaviors such as language and speech must derive from something that was previously used for a different function. In other words, all adaptations to language and speech are really exaptations. Making a distinction between speech or language-specific and general traits therefore entails making an arbitrarily defined distinction between what is meant by specific and what is meant by general. This has caused considerable confusion in the study of language evolution, as one researcher’s language-specific trait is another researcher’s domain-general trait. In contrast, the notion of whether something has undergone selective pressure related to speech or language is well defined (although of course it is in general not easy to decide whether this is the case).

Like for many biological traits, the action of selective pressure in the case of speech cannot be observed directly. Therefore, indirect means are necessary to determine which aspects of human anatomy and cognition may have undergone selective pressure related to speech. This process generally consists of three steps. First, one needs to identify a trait that has a function in speech. Then one needs to determine whether it was different in the past. This can be done by studying fossils and through comparison of homologous traits in other species. In general, the more closely related the species the better, because this means that the latest common ancestor of that species and humans lived more recently. Therefore, most often the comparison is done with apes. Finally, it needs to be determined whether the modern trait functions better than the reconstructed ancestral trait. In the ideal case, one can even show that the trait is optimal for its function in that any small change will reduce its functionality. This final step is often the most contentious, for two reasons. First, it is not always possible to determine the precise function of a reconstructed trait, and second, it is possible that the modern version of the trait only appears better because language has adapted to the trait, rather than vice versa. After all, language evolves to adapt itself to our biology at least as much as biology adapts itself to language (as will be discussed in more detail in section 23.5 on co-evolution below).

23.3.1 Anatomical adaptations

Probably the best-known example of a possible trait that has undergone selective pressure related to speech is the position of the larynx. In adult modern humans the larynx is situated lower in the throat than in other apes and therefore this phenomenon is usually called "the descent of the larynx". In addition, the velum does not descend as far as the larynx, making it impossible for adult humans to swallow and breathe at the same time. Finally, the shape of the tongue is rounder than in apes. These adaptations have the effect that the modern human tongue has greater freedom of motion than ape tongues (and presumably than those of our latest common ancestor, who lived approximately seven million years ago). However, they also have the effect of creating a higher risk of choking on our food. Presumably, the greater freedom of motion of the tongue helps to create a larger range of speech sounds. In particular, the changed shape of the tongue allows for front-back motions of the tongue body and therefore for more control over the second formant. This extended acoustic space is taken as an indication that the modifications to the modern human vocal tract have occurred because of selective pressure related to speech.

None of this is uncontroversial, however. First of all, there appears to be some descent of the larynx even in chimpanzees (Nishimura, 2003). This means that even before the emergence of speech, there were selective pressures that favored some descent of the larynx. Secondly, a modern language probably does not need the large range of speech sounds that can be produced by the human vocal tract. Cross-linguistic studies have shown that there are modern languages that use very limited repertoires of speech sounds (with Rotokas and Pirahã the languages with the smallest ones of 11 phonemes each; Maddieson, 1984). Therefore, it is argued that modern language does not need our specialized vocal tract anatomy. Finally, all tissue involved is soft tissue, and we have no fossil remains of this. The only fossil evidence that is more or less directly related to the vocal tract consists of hyoid bones from Neanderthals (who lived approximately 50 000 years ago), *Homo heidelbergensis* (who lived approximately 400 000 years ago) and *Australopithecus afarensis* (who lived approximately three million years ago) (Alemseged et al., 2006; Arensburg et al., 1989; Capasso, Michetti, & D'Anastasio, 2008; Martínez et al., 2008). The hyoid bones in Neanderthals and *Homo heidelbergensis* are similar to those of modern humans, while those of *Australopithecus afarensis* are more similar to those of chimpanzees. D'Anastasio et al. (2013) have shown that internally, too, the Kebara (Neanderthal) hyoid bone is similar to that of modern humans, indicating that muscles acted upon it in a similar way, which in turn supports a similar use of the hyoid apparatus by Neanderthals. Unfortunately, the link between the shape of the hyoid bone and that of the vocal tract is indirect. Therefore it is difficult to establish when modern human vocal anatomy emerged, and whether it is likely that this was related to speech.

In order to understand the effect of the changed anatomy on the range of sounds that can be produced, many researchers have used computer models (e.g. Boë, Heim, Honda, & Maeda, 2002; de Boer, 2010; Lieberman, Crelin, & Klatt, 1972). These do appear to indicate that the modern human anatomy has the advantage and may even be optimal. Other computer simulations appear to indicate that although languages with limited sound systems are possible, in general cultural evolution will cause sound systems to expand to fill the available articulatory space as much as possible (de Boer, 2000; Liljencrants & Lindblom, 1972; Oudeyer, 2005). In combination with the observation that the human vocal tract configuration has a distinctive disadvantage (the risk of choking on one's food, which logically needs to be compensated for with some kind of advantage) and without a convincing alternative advantage, the most parsimonious conclusion for the time being is that the vocal tract has undergone evolutionary pressure related to speech.

Another difference between humans and other apes is that other apes have air sacs, while humans do not. Air sacs in apes are large pouches with flexible walls that sit under the chin and on the chest and that are connected to the vocal tract through a tube that connects just above the vocal folds (Hayama, 1970; Hewitt, MacLarnon, & Jones, 2002). The disappearance of air sacs in modern humans and their connection to the vocal tract are an indication that their disappearance may have happened under selective pressure due to speech. Indeed, physical and computational models have shown that air sacs reduce the acoustic difference between different articulations produced with the vocal tract to which they are connected (de Boer, 2009). Experiments with human listeners have borne out that differently articulated speech sounds produced with an air sac are less easily distinguished from each other than the same articulations produced without an air sac (de Boer, 2012).

Interestingly, the shape of the hyoid bone appears to be correlated with the presence or absence of air sacs. In most apes with an air sac, the front of the hyoid bone (the basihyoid) has a cup-shaped bowl (the hyoid bulla) attached, presumably to keep the connection with the air sac open. In humans the hyoid bulla is absent. Fossil hyoid bones of Neanderthals and *Homo heidelbergensis* also lack the bulla, while the hyoid bone of an *Australopithecus afarensis* does have the bulla. From this it can be tentatively concluded that the air sac was still present in Australopithecines, but had disappeared with *Homo heidelbergensis*, the latest common ancestor of humans and Neanderthals.

However, this interpretation is not without its problems either. First of all, orangutans form an exception to the correlation of presence of bullae with presence of air sacs: they have very large air sacs, but a hyoid bone that is different from both human and other apes' hyoid bones and that does not have a bulla. This may correlate with orangutans' different morphology below the chin, and most likely consists of a development unique to the orangutan line, as other apes including gibbons do show the correlation.

Still, different alternative explanations for the disappearance of air sacs exist. In apes they may serve to produce low-frequency calls that exaggerate size, or that carry far in foliage-rich jungle. It has also been suggested that they serve to prevent hyperventilation for long, loud calls. Size exaggeration in humans may be less necessary than in other apes, because of stronger monogamy and lesser sexual dimorphism (and besides, the secondary descent of the larynx may serve to exaggerate size to some extent in adult male humans). Long-distance low calls may have been less necessary in the savannah habitat that early humans inhabited, and better breathing control (see below) may have prevented hyperventilation in extended utterances. Of course, none of these explanations are mutually exclusive, and it is entirely possible that several of these factors played a role.

Breathing control in modern humans appears to be much better than in other apes (Hewitt, MacLarnon, & Jones, 2002; MacLarnon & Hewitt, 1999). Humans can produce long, carefully controlled outbreaths during speech, while the inbreath can be very short. In ordinary speech, the outbreath is used for producing phonation, while the inbreath is silent. Other apes, in contrast, produce inbreaths and outbreaths of approximately equal duration, which tends to be much shorter on average than a human outbreath during speech. Phonation tends to be produced on both the inbreath and the outbreath. Such careful control over breathing requires precise control over the intercostal muscles. These are involved in the exhalation of air during speech.

MacLarnon and Hewitt (1999) proposed that for better control the intercostal nerves must have more fibers, and that this should be observable through a larger size of the thoracic vertebral canal (the hole in the spine through which these nerves pass). Comparison with humans and other apes have shown that the thoracic vertebral canal in modern humans is indeed larger, even when controlling for body size.

As vertebrae are often encountered as fossils, this is a promising indicator of increased breathing control in our evolutionary ancestors. Comparing the thoracic vertebral canals of a large number of fossil vertebrae, MacLarnon and Hewitt found that Neanderthals have thoracic vertebral canals of similar size to those of modern humans, while earlier hominins (*Homo ergaster* and earlier) have thoracic vertebral canals with sizes that are comparable to those of chimpanzees (always controlling for body size). This evidence indicates that breathing control appears to have been similar in Neanderthals and modern humans, while earlier hominins had breathing control similar to that of modern chimpanzees. Of course there may be other reasons for the evolution of better breathing control (long-distance running or an aquatic lifestyle have been proposed), but MacLarnon and Hewitt's argument is accepted as one of the more convincing fossil indications for speech (e.g. Fitch, 2010).

Tongue control has been investigated in a way similar to breathing control. Precise control over the tongue is necessary for speech, and it was postulated that modern humans may therefore have a larger hypoglossal nerve than other apes. This nerve emerges from the skull through the hypoglossal canal. An initial study of the hypoglossal canal showed that it is significantly larger in modern humans and Neanderthals than in late *Homo erectus*, *Homo heidelbergensis*, chimpanzees and other apes. However, a subsequent analysis of a larger sample has shown that this result does not hold up and that in addition there appears to be no usable correlation between the size of the hypoglossal nerve and the size of the hypoglossal canal. Furthermore, it turns out that the degree of tongue control involved in feeding is of the same complexity as that involved in speech (Hiiemae & Palmer, 2003). The study of tongue control therefore serves as a cautionary tale that one has to be extremely careful in studying and interpreting fossil and comparative evidence for speech.

As to the question of perception, Martínez et al. (2004) have tried to reconstruct the sensitivity for different frequencies of sound of modern humans, chimpanzees and a number of *Homo heidelbergensis* fossils. The sensitivity was reconstructed by looking at the frequency response of the middle and external ear and at the properties of the ear ossicles. They have found that modern humans have a peak sensitivity between 2–4KHz (they argue that this is an important range for speech) but that chimpanzees have diminished sensitivity in this region. The reconstructed sensitivities of their *Homo heidelbergensis* fossils also show peaks in this range, and therefore they conclude that this indicates a potential adaptation related to speech.

A potential problem with this approach, however, is that a communication system will tend to adapt itself to the capacities of its users through cultural evolution. Therefore it is an open question whether the frequency range investigated by Martínez et al. is relevant for speech because the anatomy has evolved (biologically) to adapt to speech or whether speech has evolved (culturally) to adapt to anatomy.

A final important difference between modern humans about which relatively little is known is the anatomy of the vocal folds. The precise control over phonation in modern humans is not just possible because of fine control over breathing; fine control over the precise tension and positioning of the vocal folds is also necessary (Titze, 2000). Unfortunately very little recent descriptive and comparative work has been done on vocal folds of other apes. It does appear that chimpanzee vocal fold anatomy is rather different from that of modern humans, especially in that chimpanzees have a "vocal lip" (Demolin & Delvaux, 2006; Kelemen, 1969), a small vertical extension of the vocal folds that may be involved in the very high-pitched screams that chimpanzees can produce. Unfortunately, vocal folds are small and soft structures that do not fossilize well. For the time being, the role of vocal folds in the evolution of speech will therefore remain elusive. However, it is conceivable that with a better understanding of the genetics of vocal fold development, it may be possible to achieve insights from ancient DNA. This would, incidentally, be true for all other adaptations that have been proposed for speech.

23.3.2 Neural and cognitive adaptations

The human brain is much larger than that of other apes (controlling for body size), and it has structures that appear to be specialized for language and speech. Structures in chimpanzee brains that are homologous to the ones involved with speech and language production in humans may also be involved in communication (Taglialatela, Russell, Schaeffer, & Hopkins, 2008). It appears that these structures have been exapted for language in modern human brains, but it is a subject of very lively discussion whether they have become larger due to language- and speech-specific selection or not.

The increase in brain size can clearly be seen in the fossil record, and it appears that brains tended to be as large as they are in other apes (relative to body size) in Australopithecines, but started to grow in size when the genus Homo appeared. What the role of language was in this development cannot be determined. It is possible, however, to investigate the impressions of the gyri and sulci on the inside of fossil skulls, and from this some tentative conclusions can be drawn about the development of the size of different brain regions (Schoenemann, 2006). However, it is very difficult to draw strong conclusions about the presence or absence of speech from this evidence.

Another neural adaptation to speech appears to be voluntary control over vocalization, and this has been proposed as an important distinction between humans and other apes

(Ackermann, Hage, & Ziegler, 2014; Fitch, 2010). Not only do apes have less precise control over their vocal folds than humans, but also they appear to have no or very limited voluntary control over their vocalizations. It has been proposed that a small change in the neural circuitry controlling the vocal folds has given modern humans voluntary control over their vocal folds (Fitch, 2010 calls this the Kuypers-Jürgens hypothesis). However, observations of at least one female gorilla (Koko) appear to show that she has some degree of voluntary control over her vocal folds (de Boer & Perlman, 2014). This seems to indicate that voluntary control over the vocal folds differs in degree rather than in kind between humans and (at least) gorillas. This is an area where more detailed comparative data is needed. In addition, if there really is only a simple difference in structure between the neural circuits in humans and other apes, it is possible that the genetics of it is relatively simple as well and may be traceable in ancient DNA.

23.4 Cultural evolution and sound change

In addition to the perspective of biological evolution, the perspective of cultural evolution is also useful to study phonology. There are three main reasons why this perspective is useful: first of all, the evolutionary perspective provides a systematic way to break sound change up into the processes that make up evolution: transmission, variation and selection. This in turn makes it possible to study these essentially synchronous processes in a controlled way, making it possible to better understand the essentially diachronic process of sound change. The second reason is that by reformulating sound change in terms of an evolutionary process, tools for reconstructing history that have been developed in biology can be brought to bear on questions of language history. The third reason is that a better understanding of cultural evolution helps to elucidate which aspects of language are due to historical processes operating under selective pressure for learnability and communication and which aspects of language are due to innate cognitive biases and mechanisms.

23.4.1 Systematic analysis

Blevins' (2004) evolutionary phonology is an attempt to understand and investigate the process of sound change using an evolutionary approach. She distinguishes three mechanisms underlying sound change, which she labels "change", "choice" and "chance". However, these three mechanisms do not correspond to transmission, variation and selection, but correspond to different ways in which transmission, variation and selection come about in sound change. Blevins' "change" corresponds roughly to variation due to mishearing, "choice" to variation due to mispronunciation or (possibly deliberate) simplification in pronunciation, while "chance" is variation due to ambiguity of phonological representations. Presumably transmission is implicit in Blevins' formulation, and selection is mostly due to which variants are perceived most frequently or to which patterns best fit already formed phonological representations. It should be noted that Blevins tries to avoid conscious alterations by speakers or conscious choices by learners to retain certain variants, as she appears to assume that variation must be random. However, as has been pointed out above, in cultural evolution variation and selection do not need to be random at all.

The choice for focusing on the three mechanisms of "change", "choice" and "chance" is partly a pragmatic one. These mechanisms are easier to translate into experimentally observable individual behaviors than the more abstract mechanisms of transmission, variation and selection are. They focus on perception, production and (cognitive) representation, respectively.

Blevins' aim is to explain historical sound change through the action of processes due to individual behavior. These individual behaviors can then be studied empirically, and predictions about sound change (at a time scale that is much longer than that of individual learning) can then be made by making use of knowledge of how evolution works. Conversely, it would also be possible to deduce individual behaviors from observed sound changes. In this way Blevins attempts to transform the study of sound change from a purely historical endeavor to an empirical science, just as insights into biological evolution transformed the study of biology from a mostly descriptive to an explanatory science. A basic example (Blevins, 2004, section 5.3) is Blevins' explanation of why, cross-linguistically, place features of consonant clusters tend to assimilate regressively (that is VC_1C_2V tends to become VC_2C_2V). She notes that this happens because individual perception is influenced more strongly by cues from the CV-transition than from the VC-transition. If contrasts are misidentified, this tends to happen to the first consonant more often than to the second in the cluster. "Change" therefore happens preferentially in the direction of regressive assimilation.

Blevins' approach is illustrative of modern approaches to sound change (and other types of language change, for instance grammaticalization). These approaches all try to understand properties of language as the result of the interaction between individual behaviors and cultural and historical processes. However, Blevins puts her research agenda most strongly in terms of evolutionary notions.

23.4.2 Biological tools

When formulating (cultural) language change in (biologically inspired) evolutionary terms, it becomes possible to apply to it powerful computational techniques that have been developed for reconstructing biological evolution. These tools include methods to reconstruct plausible family trees for groups of languages, methods to construct networks that represent relatedness and methods to reconstruct possible patterns of dispersal for language. Each of these methods has their origin in biology, either in the study of relatedness of DNA sequences, or in the study of the dispersal of diseases.

These techniques have mostly been applied to lexicons or to large samples of typological features, and the potential for applying them to phonological questions has not been fully exploited yet. An example from the study of word order universals may serve to illustrate how these techniques can be applied to studies of sound change. Dunn, Greenhill, Levinson and Gray (2011) used phylogenetic trees for different language families (constructed using cognate word lists) to test how well different models describing the historical change of word order fit the observed data. They found that the best-fitting models were different for different lineages and concluded that apparently word order changes depend on cultural evolution (which might behave differently in different language families), rather than on biologically determined cognitive factors (which would be the same for all language families studied). This approach could be adapted to the study of models of sound change and whether these are culture-dependent.

Building a model to reconstruct historical sound change requires making many choices about the dynamics of sound change, for instance about how likely it is that a sound changes in transmission from one generation to the next. However, it may not be sufficient to choose one average value (based on observations of well-attested historical sound changes) as such parameters may have very different values under different social and demographic circumstances. Moreover, in building a model, decisions must also be made about which mechanisms to include and which to exclude, for instance about whether to model migration, multilingualism or literacy. Therefore, although the calculations deriving probabilities for certain reconstructions are precise, the process of building the model involves many relatively arbitrary decisions. When reading papers about this topic, those decisions should therefore be carefully evaluated.

Although (as has been pointed out above) the cultural evolution of phonology is perhaps more similar to biological evolution than other aspects of language, standard techniques from biology are not always straightforward to apply to phonology for three reasons: first, biological models tend to deal with discrete traits or discrete units of DNA, whereas speech sounds are acoustic signals that can vary continuously. The second reason is related to this in that these approaches often work on the basis of the absence or presence of a trait (e.g. does the language have fricatives or not?). This often leads to very strong underuse of the available descriptive data and to disastrous results (for an example, see Atkinson, 2011; and a response by Hunley, Bowern, & Healy, 2012). The third reason is that biological models often require a model that describes evolution (e.g. voiced plosives tend to lose their voicing at the end of a word), whereas students of historical linguistics are interested in deriving the rule. All these problems can be addressed: models can be adapted to deal with continuous variation, data sets can be used in more detail and rules of sound change can be derived by comparing the performance of models that do or do not have that rule. However, this requires a lot of work, both from the builders of these models and from linguists providing the data.

23.4.3 Reducing biology

Due to selection, cultural evolution does not result in a random walk through possible phonological systems (or languages). Systems that are easy to transmit, i.e. that are easier to produce, perceive and to learn, have the advantage and therefore tend to undergo positive selection. This is in contrast to random change, which would only result in a random walk through possible phonological systems. However, if transmission is sufficiently accurate, it is even possible that (cultural) evolution results in accumulation of complexity (an example will be discussed below). Therefore, it is possible that certain complex patterns that can be observed in modern human languages are not so much reflections of cognitive constraints, but that these are the effect of cultural evolution.

Traditional accounts of phonology, from structural through generative to optimality theory-based approaches, have proposed strong innate biases or constraints as the reason behind observed “universal” patterns in languages. An example would be the preference for certain speech sounds (such as /m/ and /a/) and combinations of speech sounds (such as many languages having /s/ when they also have /z/) over others. Traditional approaches would consider such sounds and combinations “unmarked” or associated with low-ranked constraints. Although it is usually not specified explicitly, it can be assumed that these constraints must have evolved biologically under selective pressure related to language. From the perspective of cultural evolution, however, such patterns may be the result of a (historical) cultural evolutionary process under much more general functional constraints. In fact, making explanations of universal tendencies in sound systems less dependent on cognitive, innate and language-specific mechanisms, but rather analyzing them as the result of historical processes under functional constraints, was one of the aims of Blevins’ evolutionary phonology (Blevins, 2004).

An example is the emergence of combinatorial structure. Zuidema and de Boer (2009) have investigated how the use of combinatorial structure – the use of a limited number of building blocks – could become established in a population. If we have cognitive adaptations for using combinatorial structure, which is likely (Verhoef, Kirby, & de Boer, 2014), they must have evolved since the latest common ancestor with apes, as apes do not have the ability to learn and use combinatorial structure. Zuidema and de Boer showed, using a computer model based on Liljencrants and Lindblom's (1972) model for vowels, that when a system of signals is optimized for distinctiveness, the signals will begin to appear to have combinatorial structure, even though this structure is not represented explicitly in the model. However, now that the structure is present, biological, cognitive adaptations to using this structure can become advantageous. This model was also implemented as an agent-based model in which a system of signals used in a population of agents culturally evolves (de Boer & Zuidema, 2010). In this model, too, combinatorial structure emerged. Once combinatorial utterances (rather than holistic ones) become used productively in a population, it becomes possible to systematically make more complex syllables by combining the building blocks in more and more complex ways – by building complex onsets and codas, or by combining them into complex phonemes for instance – and this would be an instance of how cultural evolution can evolve towards more complexity (however, this has not been shown in models, yet).

Cultural evolution in a population of agents that all use the same rules of selection can lead to self-organization, and this is what happens not only in the above-mentioned agent-based model, but also in sound change in reality. Self-organization happens when interactions on a local scale (for instance exchanges between two speakers of a language) cause organization on a global scale (in this instance on the level of the language community). For self-organization to occur there must be positive feedback between the local and the global levels: individuals must be able to sense the global level or in any case be able to derive it from the behavior of individuals around them, and then adapt their own behavior to conform to the global level. Self-organization occurs in many living and non-living systems, and notably in language. It is not quite the same as cultural evolution: cultural evolution alone could in principle result in ever more diversity in a language community. However, part of the selective pressure in language is to conform to the perceived norms in the language community. This causes language to self-organize towards uniformity, and this is a powerful factor in cultural evolution of language. Of course, when the language community becomes too large or too dispersed for speakers to sense the norms at the global level, the community may split into different dialects.

There are three reasons why it is essential to reduce the number of cognitive adaptations that need to be explained by biological evolution. First of all, there is emerging insight that cognitive adaptations to a rapidly changing culturally transmitted system can only evolve under certain rather specific circumstances (this will be discussed in the next section). Secondly, there has been relatively little time for biological adaptations to evolve. The little fossil evidence that exists appears to indicate that the first adaptations to speech must have evolved somewhere between *Homo erectus* and *Homo heidelbergenis*. This entails a time span of about a million years, and this is actually relatively short for species with a long generation gap, such as apes. Finally, from a methodological perspective it is more satisfactory to have an explanation based on pre-existing factors without the need of adding something extra (a cognitive adaptation) to the explanation. This does not mean that there are no (cognitive) adaptations to learning and using speech. However, in accounts of language that make use of cultural evolution they are probably fewer in number and more general than the mechanisms (such as distinctive features and their markedness) that have been proposed in non-evolutionary theories.

23.5 Interaction between cultural and biological evolution

Both biological evolution and cultural evolution play a role in determining what a language and its phonology look like. The interaction between these two evolutionary processes makes it more complicated to study cognitive adaptations related to speech and language ("the human capacity for language"). Cultural evolution generally operates much more quickly than biological evolution – sound systems change over the order of hundreds of years, while other aspects of language can change over the order of years or decades. Biological change generally needs thousands of years to operate. This means that language and speech may be a moving target for biological evolution, making it harder for adaptations for language to evolve. At the same time, cultural evolution and self-organization may cause language and speech to show universal tendencies that may appear (and have been analyzed in the past) as being the result of cognitive mechanisms. The interaction between biological and cultural evolution therefore raises two questions: what kind of biological adaptations to something as ephemeral as speech and language *can* evolve and how do we tell the effect of (biologically evolved) cognitive mechanisms from the effect of cultural evolution when studying language?

23.5.1 What kind of adaptations can evolve?

In the case of adaptations for learning language, cultural evolution must drive biological evolution. This is because (linguistic) communication by definition involves a sender and a receiver, and it would not help if one of them has a biological innovation (to improve learning of language) that the other members of the population cannot deal with. For biological innovations to confer an advantage, they must therefore help to deal with something that is already present in the population. In the case of language and speech, whatever is present must be due to cultural transmission and evolution. Thus evolution of cognitive adaptations for learning consists of originally entirely general learning behaviors becoming more and more specialized for speech and language. This process is called the Baldwin effect (Baldwin, 1896).

The Baldwin effect is a very general mechanism by which an originally acquired trait becomes gradually more and more innately (genetically) specified because there is an advantage in acquiring the trait as quickly as possible. Therefore slight mutations that help to acquire or that already pre-code part of the trait will have an evolutionary advantage. Over time an originally acquired trait can become entirely innate. Pinker and Bloom (1990) proposed that language, or more specifically universal grammar, has evolved in this way.

However, the Baldwin effect only works when the environment in which it operates is stable over time. This is not the case for language because it evolves culturally much more rapidly than biology. Computer simulations of this interaction (Chater, Reali, & Christiansen, 2009) indicate that because of this, biological adaptations to arbitrary, non-functional aspects of language cannot evolve biologically. Therefore, Chater, Reali and Christiansen conclude that functionally arbitrary principles, features or markedness constraints cannot be innately specified due to language-specific mechanisms. Accordingly, they propose any such phenomena observed in language must be due to domain-general mechanisms and cultural evolution. Of course, this leaves open the possibility of the evolution of adaptations to language that improve its function, and to properties of language that are stable over time.

Interestingly, cultural evolution may create such stability, as it tends to move linguistic systems towards certain preferred configurations. Those configurations will always be in some sense optimal, as cultural evolution of language tends to optimize its transferability from generation to generation. An example would be the maximal use of the available articulatory space. Computer simulations (de Boer, 2000; Liljencrants & Lindblom, 1972; Oudeyer, 2005) have shown that systems of speech sounds will tend to use the available articulatory space maximally. This in turn puts pressure on biological evolution to evolve ever-larger articulatory spaces. Another example could be the example of evolution of combinatorial structure that was mentioned in the previous section. Outside the domain of phonology, the evolution of lexicon size could be an example: even though the words in a language change relatively rapidly, the size of the lexicon remains constant over time (de Boer, 2014).

23.5.2 The relation between language and cognitive mechanisms

The interaction between cultural evolution and biological evolution not only calls into question what kind of adaptations to language and speech can evolve; it also makes it more difficult to determine whether an observed property of language is due to biologically specified cognitive mechanisms or whether it is an effect of cultural evolution. The simulation of the emergence of combinatorial structure discussed in the section on cultural evolution serves to illustrate this: although in that simulation there were no mechanisms whatsoever for dealing with combinatorial structure, nevertheless it emerged in the simulations, purely as the result of cultural evolution. Therefore one should be careful not to conclude from observing a linguistic phenomenon that there must necessarily be cognitive mechanisms that underlie that phenomenon. An observed linguistic phenomenon may or may not reflect an underlying cognitive mechanism.

Understanding the relation between cognitive mechanisms and cultural evolution has been the goal of a number of mathematical and computational modeling efforts. These models are based on a population of agents that can each learn a simple language. This language can take a number of forms, and although each agent can learn each language type, it has a preference for certain types over others. These are called learning biases, and they correspond to certain types of language (or sound systems, or whatever is learned) being more readily acquired than other types. Agents not only learn a language, they also reproduce it, with a certain amount of random variation. Their output is then learned by the next generation of agents. This process is called iterated learning (Smith, Kirby, & Brighton, 2003). The mathematical analysis then investigates what happens in the long run: what form does the language eventually take given learning biases of the individual agents?

Griffiths and Kalish (2007) have analyzed the case of Bayesian learners: a type of idealized rational agent that uses the linguistic data it observes together with its own learning biases to calculate the probability of each possible type of language. When the variation in the linguistic (re-)productions of such Bayesian agents is exactly the variation they have observed during learning (this is called probability matching), then the languages that eventually emerge exactly reflect the biases of the agents. That is, the frequency with which different variants of the language occur exactly correspond to the strengths of the biases that the individual agents have. In this case, observation of linguistic patterns would therefore tell us much about the cognitive mechanisms of the language users.

However, Kirby, Dowman and Griffiths (2007) have shown that this result critically depends on how agents reproduce variation. If the agents do not exactly reproduce the variation they observe but instead tend to preferably produce the most frequently observed variants (such agents are called maximizers), the relation between learning biases and the frequency of language variants becomes more complex. Even very weak biases in favor of certain variants cause those variants to come to dominate the eventual language that emerges. Therefore, in the case that agents are maximizers, it is not possible to draw precise conclusions about agents' learning biases by observing the distribution of language variants. Unfortunately it is as yet unknown whether human language learners behave more like probability matchers or like maximizers. It is even possible that humans tend to behave differently at different stages of linguistic development and for different aspects of language. If we are to understand the precise relation between typological observations and learning mechanisms, we need to learn more about this aspect of human linguistic behavior.

23.5.3 Experimental studies

Questions relating to evolution of language and to the interaction between biology and culture are now studied experimentally, using variations on the experimental iterated learning paradigm (Scott-Phillips & Kirby, 2010). As the terminology has not yet crystalized, the terms experimental semiotics and experimental cultural learning are also used. This paradigm was originally used to investigate the emergence of simple communication systems (Galantucci, 2005) and to investigate the emergence of structure (Kirby, Cornish, & Smith, 2008). In iterated learning experiments, participants are required to learn, reproduce and often generalize language-like items. However, the tasks are generally formulated in such a way that participants cannot use the linguistic knowledge they already have. There are two main variants: in one variant, participants are required to develop a communication system from scratch in cooperation with one or more other participants (Galantucci, 2009). This models emergence of conventions in a population. In the other variant, participants are presented with a number of language-like items (signals or form-meaning mappings) that they need to learn and later reproduce. The next participant in the experiment is then trained on the basis of these utterances (Scott-Phillips & Kirby, 2010). This models language transfer between generations. Importantly, participants engage not just in learning but also in active reproduction, whereas in more classical psycholinguistic experiments participants are tested on passive knowledge after having been exposed to training stimuli. However, it is possible to design intermediate experiments in which participants are trained on a precisely defined set of stimuli and are then asked to reproduce these. Many experiments in this experimental paradigm do not deal with complex signals at all and use symbolic strings (often consisting of characters read from a computer screen and typed with a keyboard) as their stimuli. However, there are a growing number that are of direct relevance to the study of phonetics and phonology.

Iterated learning experiments tend to focus on questions related to the origins and emergence of language, rather than on questions related to language change. Experiments concerned with the emergence of signals tend to focus on the emergence of combinatorial structure and on the tradeoff between combinatorial structure and iconic structure. This is an important question in the study of the evolution of language and speech, as there is much speculation about whether initial stages of language used iconic structure. When there is iconic structure – signals resemble what they refer to – signals are easier to understand and learn. However, iconic structure would block the systematic reuse of building blocks – combinatorial structure – that human languages tend to have. This has been observed in the emergence of Al-Sayyid Bedouin Sign Language (Sandler, Aronoff, Meir, & Padden, 2011). This sign language, which has emerged over the last 70 years, apparently has very little combinatorial structure, but more iconic structure. The questions of how the tradeoff between combinatorial and iconic structure works in the emergence of language, how it is influenced by social circumstances and what is the contribution of cultural processes and of cognitive mechanisms, are investigated by experimental iterated learning studies.

The first experiments (Galantucci, 2005) on emergence of systems of signals used a device that was designed to produce continuous visual signals and that could not be used to simply draw or write text. Galantucci looked at the structure of the emerging signals and at what kinds of signals were more successful in communication. However, this analysis was only a sideline of these experiments. Fay, Garod, and Roberts (2008) have looked specifically at the properties of systems of visual signals (small drawings) that emerge in a communication task. When there is repeated interaction between only two participants, signals quickly become simpler, and iconic structure disappears. When the group consisted of eight participants, however, signals tended to stay more iconic. More recently, Roberts and Galantucci (2012) have explored the emergence of structure in these visual stimuli more in depth and looked at the tradeoff between iconic structure and combinatorial structure. It turns out that combinatorial structure does not emerge as readily when participants can use iconic structure. The effect of different signaling spaces on this tradeoff is a current topic of research (Little & de Boer, 2014; Roberts & Galantucci, 2014). The effect of memory constraints on structure and simplification is being investigated by Tamariz and Kirby (2014).

Del Giudice (2012) and Verhoef, Kirby and de Boer (2014) have looked explicitly at the question of whether combinatorial structure is the result of self-organization and cultural evolution or whether modern humans have cognitive adaptations for detecting and generalizing structure. Verhoef et al.'s experiments have shown that modern humans do have cognitive mechanisms for dealing with combinatorial structure. Although self-organization can explain the initial emergence of combinatorial structure, and it was therefore not strictly necessary early in evolution to have cognitive mechanisms for dealing with it, modern humans represent and use combinatorial structure explicitly.

Using a more abstract approach, Ferdinand, Kirby and Smith (2014) have investigated how participants generalize probabilities, in order to investigate whether humans are probability matchers or maximizers. Their experiments are based on drawing different colored balls from urns, and participants are tested in different ways what probabilities for the different colors they learn on the basis of their observations. It turns out that humans are neither maximizers nor probability matchers, but that they show intermediate behavior that depends on the context. Their work is being adapted to continuous signals by van der Ham and de Boer (2014). This work again shows that (at least adult) participants do not engage in simple probability matching. These observations make it less likely that there is a simple mapping between observed patterns of language phenomena and the cognitive mechanisms underlying them.

The study of the evolution of phonology is a relatively new field that has recently grown along with the renaissance of the scientific study of language evolution. Many of the findings, questions and methods discussed above are still in full development, and many contributions can still be made. The above overview is therefore as much a guide to the open questions and possible avenues of research as it is an overview of existing results.

23.6 Background reading

There are many good textbooks on evolution, for instance Herron and Freeman (2014) or Futuyma (2013). For a more popular introduction to the complexities of evolution, *The Selfish Gene* by Richard Dawkins is, although a bit older, still a good choice. For more information about the mathematics of evolution, Nowak (2006) is an accessible source and contains a chapter about language evolution. It is, however, very much centered on Nowak's own research.

A thorough introduction to language evolution is Fitch (2010). A more accessible brief introduction is Hurford (2014). Aiello and Dean (2002) provide a good introduction to the anatomical evolution of humans.

For a starting point on literature of the experimental study of evolution of language, Galantucci (2009); Scott-Phillips and Kirby (2010); and Kirby, Griffiths and Smith (2014) can be recommended.

For an introduction to cultural evolution, Richerson and Boyd (2005) is a good starting point, while Levinson and Jaisson (2006) contains a good collection of more technical essays on the topic. Both are about much more than language alone.

References

Ackermann, H., S. R. Hage & W. Ziegler. 2014. Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences*, *37*(6), 529–546.

Aiello, L. C. & C. Dean. 2002. *Human Evolutionary Anatomy*. Amsterdam: Elsevier Academic Press.

Alemseged, Z., F. Spoor, W. H. Kimbel, R. Bobe, D. Geraads, D. Reed & J. G. Wynn. 2006. A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature*, *443*(7109), 296–301.

Arensburg, B., A. M. Tillier, B. Vandermeersch, H. Duday, L. A. Schepartz & Y. Rak. 1989. A Middle Palaeolithic human hyoid bone. *Nature*, *338*(6218), 758–760.

Atkinson, Q. D. 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, *332*, 346–349.

Baldwin, J. M. 1896. A new factor in evolution. *The American Naturalist*, *30*, 441–451, 536–553.

Blevins, J. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boë, L.-J., J.-L. Heim, K. Honda & S. Maeda. 2002. The potential Neanderthal vowel space was as large as that of modern humans. *Journal of Phonetics*, *30*(3), 465–484.

Capasso, L., E. Michetti & R. D'Anastasio. 2008. A Homo erectus hyoid bone: Possible implications for the origin of the human capability for speech. *Collegium Antropologicum*, *32*(4), 1007–1011.

Chater, N., F. Reali & M. H. Christiansen. 2009. Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences*, *106*(4), 1015–1020.

D'Anastasio, R., S. Wroe, C. Tuniz, L. Mancini, D. T. Cesana, D. Dreossi & L. Capasso. 2013. Micro-biomechanics of the Kebara 2 hyoid and its implications for speech in Neanderthals. *PLoS ONE*, *8*(12), 1–7.

De Boer, B. 2000. Self organization in vowel systems. *Journal of Phonetics*, *28*(4), 441–465.

De Boer, B. 2009. Acoustic analysis of primate air sacs and their effect on vocalization. *The Journal of the Acoustical Society of America*, *126*(6), 3329–3343.

De Boer, B. 2010. Investigating the acoustic effect of the descended larynx with articulatory models. *Journal of Phonetics*, *38*(4), 679–686.

De Boer, B. 2012. Loss of air sacs improved hominin speech abilities. *Journal of Human Evolution*, *62*(1), 1–6.

De Boer, B. 2014. Biological adaptations to cultural traits. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 419–420). Hackensack, NJ: World Scientific.

De Boer, B. & M. Perlman. 2014. Physical mechanisms may be as important as brain mechanisms in evolution of speech. *Behavioral and Brain Sciences*, *37*(5), 552–553.

De Boer, B. & W. Zuidema. 2010. An agent model of combinatorial phonology. *Adaptive Behavior*, *18*(2), 141–154.

Del Giudice, A. 2012. The emergence of duality of patterning through iterated learning: Precursors to phonology in a visual lexicon. *Language and Cognition*, *4*(4), 381–418.

Demolin, D. & V. Delvaux. 2006. A comparison of the articulatory parameters involved in the production of sounds of bonobos and modern humans. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)* (pp. 67–74). New Jersey: World Scientific.

Dunn, M., S. J. Greenhill, S. C. Levinson & R. D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, *473*(7345), 79–82.

Fay, N., S. Garrod & I. Roberts. 2008. The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society of London B*, *363*, 3553–3561.

Ferdinand, V., S. Kirby & K. Smith. 2014. Regularization in language evolution: On the joint contribution of domain-specific biases and domain-general frequency learning. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 435–436). Hackensack, NJ: World Scientific.

Fitch, W. T. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.

Futuyma, D. J. 2013. *Evolution, Third Edition*. Sunderland, MA: Sinauer Associates, Inc.

Galantucci, B. 2005. An experimental study of the emergence of human communication. *Cognitive Science*, *29*, 737–767.

Galantucci, B. 2009. Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, *1*, 393–410.

Griffiths, T. L. & M. L. Kalish. 2007. Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*(3), 441–480.

Hayama, S. 1970. The Saccus laryngis in primates. *Journal of the Anthropological Society of Nippon*, *78*(4), 274–298.

Herron, J. C. & S. Freeman. 2014. *Evolutionary Analysis, Fifth Edition*. San Francisco, CA: Benjamin Cummings.

Hewitt, G. P., A. MacLarnon & K. E. Jones. 2002. The functions of laryngeal air sacs in primates: A new hypothesis. *Folia Primatologica*, *73*, 70–94.

Hiiemae, K. M. & J. B. Palmer. 2003. Tongue movements in feeding and speech. *Critical Reviews in Oral Biology & Medicine*, *14*(6), 413–429.

Hunley, K., C. Bowern & M. Healy. 2012. Rejection of a serial founder effects model of genetic and linguistic coevolution. *Proceedings of the Royal Society B*, *279*(1736), 2281–2288.

Hurford, J. R. 2014. *The Origins of Language: A Slim Guide*. Oxford: Oxford University Press.

Kelemen, G. 1969. Anatomy of the larynx and the anatomical basis of vocal performance. In G. H. Bourne (Ed.), *The Chimpanzee* (Vol. 1, pp. 165–186). Basel: Karger.

Kirby, S., H. Cornish & K. Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Kirby, S., M. Dowman & T. L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, *104*(12), 5241–5245.

Kirby, S., T. L. Griffiths & K. Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*(1), 108–114.

Levinson, S. C. & P. Jaisson. 2006. *Evolution and Culture*. Cambridge, MA: MIT Press.

Lieberman, P. H., E. S. Crelin & D. H. Klatt. 1972. Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee. *American Anthropologist*, *74*, 287–307.

Liljencrants, J. & B. Lindblom. 1972. Numerical simulations of vowel quality systems. *Language*, *48*, 839–862.

Little, H. & B. de Boer. 2014. The effect of size of articulation space on the emergence of combinatorial structure. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 479–480). Hackensack, NJ: World Scientific.

MacLarnon, A. & G. P. Hewitt. 1999. The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology*, *109*(3), 341–343.

Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
Martínez, I., J.-L. Arsuaga, R. Quam, J.-M. Carretero, A. Gracia & L. Rodríguez. 2008. Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *Journal of Human Evolution*, *54*, 118–124.
Martínez, I., M. Rosa, J.-L. Arsuaga, P. Jarabo, R. Quam, C. Lorenzo, . . . E. Carbonell. 2004. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences*, *101*(27), 9976–9981.
Nishimura, T. 2003. Comparative morphology of the hyo-laryngeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. *Primates*, *44*(1), 41–49.
Nowak, M. A. 2006. *Evolutionary Dynamics*. Cambridge, MA: Harvard University Press.
Oudeyer, P.-Y. 2005. The self-organization of speech sounds. *Journal of Theoretical Biology*, *233*(3), 435–449.
Pinker, S., & P. Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707–784.
Richerson, P. J. & R. Boyd. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: The University of Chicago Press.
Roberts, G. & B. Galantucci. 2012. The emergence of duality of patterning: Insights from the laboratory. *Language and Cognition*, *4*(4), 297–318.
Roberts, G. & B. Galantucci. 2014. The effect of iconicity on the emergence of combinatorial structure: An experimental study. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 503–504). Hackensack, NJ: World Scientific.
Sandler, W., M. Aronoff, L. Meir & C. Padden. 2011. The gradual emergence of phonological form in a new language. *Natural Language & Linguistic Theory*, *29*(2), 503–543.
Schoenemann, P. T. 2006. Evolution of the size and functional areas of the human brain. *Annual Review of Anthropology*, *35*, 379–406.
Scott-Phillips, T. C. & S. Kirby. 2010. Language evolution in the laboratory. *Trends in Cognitive Sciences*, *14*, 411–417.
Smith, K., S. Kirby & H. Brighton. 2003. Iterated learning: A framework for the emergence of language. *Artificial Life*, *9*(4), 371–386.
Taglialatela, J. P., J. L. Russell, J. A. Schaeffer & W. D. Hopkins. 2008. Communicative signaling activates "Broca"s' homolog in chimpanzees. *Current Biology*, *18*(5), 343–348.
Tamariz, M. & S. Kirby. 2014. Culture: Copying, compression and conventionality. *Cognitive Science*, *31*(1), 171–183. DOI: 10.1111/cogs.12144.
Titze, I. R. 2000. *Principles of Voice Production*. Iowa City, IA: National Center for Voice and Speech.
Van der Ham, S. & B. de Boer. 2014. The effect of iconicity on the emergence of combinatorial structure: An experimental study. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 541–542). Hackensack, NJ: World Scientific.
Verhoef, T., S. Kirby & B. de Boer. 2014. Emergence of combinatorial structure and economy through iterated learning. *Journal of Phonetics*, *43*(C), 57–68.
Zuidema, W. & B. de Boer. 2009. The evolution of combinatorial phonology. *Journal of Phonetics*, *37*(2), 125–144.

Language index

Subject index